

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 02:48:50 ; Search time 5973 Seconds

(without alignments)
11292.426 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaataaccatcatgatac.....ccggcggaagtcgtcgat 1392

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hvg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sbs: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	628.8	45.2	11493	1	AF283498	AF283498 Azospirillum
C 2	555.6	39.9	300883	1	AE016809	AE016809 Vibrio vulnificans
C 3	541.2	38.9	248850	1	AP005348	AP005348 Vibrio vulnificans
C 4	516.2	37.1	302100	1	AP005087	AP005087 Vibrio parvulus
C 5	431	31.0	311050	1	BX294133	BX294133 Pirella
C 6	171.2	12.3	110000	1	AE017282_06	Continuation (7 of
C 7	133	9.6	10676	1	AE008293	AE008293 Agrobacterium
C 8	133	9.6	10707	1	AE009321	AE009321 Agrobacterium
C 9	122.8	8.8	348077	1	AP003000	AP003000 Mesorhizobium
C 10	119.8	8.6	300150	1	AP005945	AP005945 Bradyrhizobium
C 11	118.2	8.5	110000	1	RME591985_07	Continuation (8 of
C 12	116	8.3	196050	1	AL646058	AL646058 Ralstonia
C 13	113.2	8.1	298700	1	AP005951	AP005951 Bradyrhizobium
C 14	109.6	7.9	12129	1	AE007223	AE007223 Sinorhizobium
C 15	104.8	7.5	299750	1	AP006575	AP006575 Gloeobacter
C 16	104.4	7.5	339650	1	AP003583	AP003583 Nostoc sp
C 17	100.6	7.2	348068	1	BX572604	BX572604 Rhodospirillum
C 18	98.4	7.1	300700	1	AP006573	AP006573 Gloeobacter
C 19	98.2	7.1	110000	1	BX950851_33	Continuation (34 of

20	95.2	6.8	17155	1	AY273169	AY273169 Rhodospirillum rubrum
21	95	6.8	4148	1	KPNASREC	K2741 Klebsiella pneumoniae
C 22	93.4	6.7	13551	1	D90906	D90906 Synecococcus
C 23	89.8	6.5	3253	1	SYONIRNRT	D12723 Synecococcus
C 24	89.2	6.4	300363	1	AE016781	AE016781 Pseudomonas
C 25	87	6.2	10798	1	AE004604	AE004604 Pseudomonas
C 26	86	6.2	300300	1	AP005373	AP005373 Rhodospirillum rubrum
27	84.6	6.1	349142	1	BX572599	BX572599 Rhodospirillum rubrum
28	82.4	5.9	1371	6	AR36463	AR36463 Sequence
C 29	82.4	5.9	299650	1	AP005955	AP005955 Bradyrhizobium
C 30	80.2	5.8	4080	1	SYOCMP	D26358 Synecococcus
C 31	79.4	5.7	298750	1	AP005375	AP005375 Rhodospirillum rubrum
C 32	77.8	5.6	302550	1	AP006581	AP006581 Gloeobacter
C 33	75.2	5.4	11318	1	AE008005	AE008005 Agrobacterium
C 34	75.2	5.4	11330	1	AE009039	AE009039 Agrobacterium
C 35	74.6	5.4	6449	1	PLNRTABC	Z1958 P. lamiosum
C 36	74.4	5.3	347400	1	AP003591	AP003591 Nostoc sp
C 37	73.8	5.3	7520	1	AE005736	AE005736 Caulobacter
C 38	72	5.2	132106	1	SYCSLTH	D64006 Pseudomonas
C 39	71.8	5.2	310325	1	AE016864	AE016864 Pseudomonas
C 40	70	5.0	3218	1	AF001333	AF001333 Synecococcus
C 41	69.2	5.0	11321	1	AE005737	AE005737 Caulobacter
C 42	65.8	4.7	3719	1	AY283674	AY283674 Cyanobacter
C 43	63.4	4.6	2210	1	AVNASST	X83602 A. vinei
C 44	63.4	4.6	4080	1	SPCCNRT	X61625 Synecococcus
C 45	60.8	4.4	110000	1	CP000011_11	Continuation (12 of

ALIGNMENTS

RESULT 1	AF283498	11493 bp	DNA	linear	BCI 02-JUL-2002
LOCUS	Azospirillum brasilense nitrate assimilation gene cluster, complete sequence.				
DEFINITION	AF283498				
ACCESSION	AF283498.1	GI:21666012			
VERSION					
KEYWORDS					
SOURCE	Azospirillum brasilense				
ORGANISM	Azospirillum brasilense				
REFERENCE	1 (bases 1 to 11493)				
AUTHORS	Steenhoudt, O., Ptacek, D., Verreth, C. and Vanderleyden, J.				
TITLE	Nitrate assimilation in Azospirillum brasilense Sp245: mutational and expression analysis of the nas gene cluster				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 11493)				
AUTHORS	Steenhoudt, O.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-JUN-2000) F.A. Janssens Laboratory of Genetics, Catholic University of Leuven, Kard. Mercierlaan 92, Heverlee 3001, Belgium				
FEATURES	Location/Qualifiers				
source	1..11493				
	/organism="Azospirillum brasilense"				
	/mol_type="genomic DNA"				
	/strain="Sp245"				
	/db_xref="taxon:192"				
gene	complement(526..2190)				
	/gene="nasD"				
CDS	complement(526..2190)				
	/gene="nasD"				
	/codon_start=1				
	/transl_table=11				
	/product="nitrate transporter AtPase"				
	/protein_id="AA073541.1"				
	/db_xref="GI:2166013"				
	/translation="MAILELKNVAKSYGTEVLRGIDLSIAEGEFAIVFGSGAKTT				
	LVNMGSLAADPAGVTLFPGAVRGGAEGVVFQVYSIMPLSYDNVALAVDAVHK				
	GKSPRGVLTBRAVETVGLGATVTRPRELGGKRGVRFAPALAMSEMLLDLPL				
	SALDLTRAKODEIEAIWRKQRTVVLITNDVDAIILADNIITLPFGATLGRF				

[illegible]

gene
CDS

/translation="MDQSIWITVGSVEDIPROGSIWVTRTFEDIALFTVVADEVYAL
EDICPHAGGPGLSGIVHGARVCTPLHSWIIDLASGAVGPDSCAGHCPIRVVDGAQQ
LGIAAFPRGGRGGCGHG"
8473. 11130
/gene="naaa"
8473. 11130
/gene="naaa"
/codon_start=1
/transl_table=11
/product="naseimlatory nitrate reductase catalytic
subunit"
subunit_id="AAM73546.1"
/protein_id="AAM73546.1"
/db_xref="GI:21666018"
/translation="METVTRPCYCGVGGCVLASVSDGTGVAGADPDHPANRKLCS
KGINLPPTTGLAERLLHPVDGRRVAMDEIAVARRLSPTLAKHPDSVAFVPSOL
LTEDYAAVANKLAKGPIGTANIIDPNSRLCAASSVAHGKRGAGDAVYITDDLEADLV
LTVGSINAMCHPTLNORLLAAKHTGRIVLIDPRRTAAVDGADRAVLPTRPGDLVLF
NGLLYVLADNHLADNGLIDARFVAERTAGLAEALEARADAPRAHVAFCGGLPLAEV
EAYVAEYAVTRRTVTVYSGVNSOGSTGTVNAIIVNHFTGRTGAPMGPFSEVCTOP
NANGGEVGLGANQLAAHGGFESVDRLRRFNNAVRVADPKLKVADPFAVEGCTP
RAYWNAITPAVAMPDAARVRRALRCAETVIVYSCVADDTATLSHILPAAQMSKD
GTYTNSDRTISRORAFRPLAGEARPDWVIYTVARAMGPAASFEETPADVFERHAL
SAEENGSTRAFIDIGLAGLSDAEFTTIAAPRRRAGEGTRRLFADGRFTVEDGAR
FVIVASRPLPRSVPEGSLMINTGRILDQPHMTWRTGSLPSLSHAPEPCIDLPDOAA
ARGLTGALVLSSTPGTALARLRVTEADRVGSAFLPMHTIDFTSLSCVGRALVDDA
VDVSPGQDILKMPVTLTPYPAAMGTGFLRAVEVEPAHAGCYSRAVPGSHITLSE
DALPAPETLAFGPAAGTVDYSDAPGVLERAWYTHVRLNACLPLETSLSGALPISGMILVE
LTSAEALDGAARALISGRSPVETADEGRWVSCFVGWNTLTAIQLISQGLTTPVQEVG
AALRAGTNGSCSIPELKEIILADQRAHVA"

ORIGIN

Query Match 45.2%; Score 628.8; DB 1; Length 11493;

Best Local Similarity 68.7%; Pred. No. 8.5e-11;
Matches 912; Conservative 0; Mismatches 407; Indels 9; Gaps 3;

69 GGGTTTGAAGATTGCGCCCGATGTGGGGCAGTTGGCAAGCTGAAAAAGAGATTGAA 128

4787 GGCGTTGATGCTCGGCTCCGCCAGGCGCTCCGCTCGACGTGGAGAAGACCAAGCTCAA 4728

129 ATTCCGGCTTCATCAAGCTCACCAGACATGGCGCCGCTGGCGGTGGCGCCCGGAAAAAGGCTT 188

b
4727 GCTGGGCTTCATCAAGCTGACCGATATGGCCCCGCTCGCCATCGCCGCGGAGAGAGGCTT 4668

189 CTTGAGAGACGAGGCTGTTCTGCACTGGAAGCGACGCCAATGGAAGTGTGAT 248

b
4667 CTTGAGGACGAGGCGTGTGGTGACGCTGAGCCGCAAGCCAACTGGAAGTGTGCT 4608

249 GGATAGGGTCGTGAAATGGCGAAGTGGCTGCACATGCTGGCGCGGCGCTTACG 308

b
4607 CGACCGTGTGATCTCCGGGAGCTGGAACGGCGGCCACATGCTGGCCGGCCAGCCGCTGGG 4548

309 GGCACGCGTTGGCTTCGGACCAAGCGCATATCGAGTGCCTTCAGCATGGGCTTCAA 368

b 4547 GGCCACCAATCGGCTTCGGACCCAGGCCAAGTGTGACCGCTTCTTCGATGACCTGAA 4488

369 CGGCACCGGATTACGGTGTCCATGAATCTGGCATCAATGAGCGGAACATACC--- 425

b

4487 CGCGAAGGCATCACTCTGTCCAAACGAGTGTGGGAGCGCATGCCCAACTGCCCCA 4428

426 GCTGGAGGCGGTAACCCGGTCCATCCGATCAGGCGAGATATCTAAACCCGGTCGTGA 485

485 A A A A T T C A A A G C C C G A A G G C A A C C C G T T T C A A T A T A T G G C G A T G C A C C T T T T C C C C C C G G A T T C T T A 545
492 / G G G G C C G G A C G C A A G C C C G C T G C A C C C G A T C A A G G C C G A C C G C T G A A G C C C G G T C A T C C 4308

4367 CCACTAATCGGGGACGAGGGGCAAGCCCTTACCAACAGGGGCAATGGTCTCTTCCCGGTCCACCCA 4308

546 CAACATCAACTGCGTTACTGGCTGGCGGCTGCGGCTATCAATCTGGCTACTATTGCC 605

4307 CAATTACGAGCTGCGCTACTGCTTGGCGCGCGCGGACATCAACCCCGGCTACTA--CGC 4251

606 GCCGCAAGACATTTCCGGCCAAATCGGCGCAGACGCCCTTGTTGTCCGTGACCCCGCGGCC 665

gene
5180. .6094
/locus_tag="VV20277"
5180. .6094
/locus_tag="VV20277"
/note="COG0583"
/codon_start=1
/transl_table=11
/product="Transcriptional regulator"
/protein_id="AAO07242.1"
/db_xref="GI:27358292"
/translation="MILITKVEVSPPTTEALHILDAIDRGSFASANEIDRAESS
LSYQIAQEDQIDIMIPDRSGKANFTAGYLIERGTTILAAETKLVNATYLLANCM
ELDIQ1ADVDIGIHHPSLDALGNVSKTRVLRLOEBILACGWSLQSGDANFVKR
LEVIDHVAADTITGKMSWALVAAPHYVAKSGSGFNDRARKEYIIAIAPTAPOPI
TVNIIOKOPRLVTNLOAKVDAVLVGLGIGTLPDIAFTPLLEAKLKEIGABQGHLE
IILAKRMQMGSAKWCIOYLKKMKLK"

gene
5180. .6094
/locus_tag="VV20278"
5180. .6094
/locus_tag="VV20278"
/note="COG0583"
/codon_start=1
/transl_table=11
/product="Histone acetyltransferase HPA2"
/protein_id="AAO07243.1"
/db_xref="GI:27358293"
/translation="MKKIQKLFKPTSTAVIGASQRHLSAGNIWNKNLQSGFOGTTIP
VTRKSVAGVAVPSVTELPVLPDIAITCRSEKNVALFKQLEIGTEFVIVASTDT
DYPHDDQSDVADACRIARIESNMRYLGNISGLILPWKFNASPTVARGKIAFTIS
QSAACCTTILIDANDKEIGFSAFVSIINGLIDIDELIDYLDSTSKTDALLVYDSTT
DARRTSAPARASRRRLIVLKGRSILGRALNKGADTLDIYVDSAIRSGMLRVN
NMHELPAVETVIRVQVIRGERTAITTGAGPALMAVDITFERGGLANISSETEML
SGHSSSMHNHPDIDVGDADKRYVATINAILDSDCADVALMHSSAYSDSVETQ
ALVDAIQSPHPRKRNILINWSGEQTAKPAMTILTKGIPYRTPESAVAPMHLVRY
RNQOQLQSEPTTAPAEPLVSDVNAKAWSEHMDKEJAOUDTQIGSLKLPIFDVL
PMTAISDTEVTHIAEQIGYPAVNLSPDIAHSDVOGVMLNLSQGEVASADAIL
DRTKLAPYANHGLVOGMAKLAGSEIRVAVITDKFGVUILGOGSGMSBARDA
TALPLPMTATVAVOIVRAIKEGHRLQKLEPMDVYGLAKETVISOMVVEIPOJKE
LDIHVPLANGEHFTILDADTLNENTLGHGQERLAIRPPAEFVETMQLKGEPIILAP
LILPEBESHAERINNVSKEBDYKRFSDSFGNEHALANLQIYDIDROMKVAASFSS
GSAKIGVAPRALINPENTAEPAILIRSDLVKGKGGKGLKLNKIIDYCRGKTKMTGM
TTPTRKMTLTLQKIGFSLDIQFEDGVADMTPLNP"

gene
9046. .9621
/locus_tag="VV20279"
9046. .9621
/locus_tag="VV20279"
/note="COG3266"
/codon_start=1
/transl_table=11
/product="Uncharacterized protein conserved in bacteria"
/protein_id="AAO07244.1"
/db_xref="GI:27358294"
/translation="MKKVAITLTVLSACASDNYVANVTTESHREKTEPIPEPIM
AOEYVQVQSIPOQVAKMEPOPEKQKVYKLPQSTTAVAVIPVPSKODKQKORFQY
TVQVAVNSQAKVADVFAKSLPDKDQKIVNENKYNIGTMTFVVLGDPATISTEAKAIA
QUPQEKELQRPVKSIDAIKNSIPTLNKLN"

gene
9788. .10783
/locus_tag="VV20280"
9788. .10783
/locus_tag="VV20280"
/note="Similar to ATP-grasp enzymes; COG1181"
/codon_start=1
/transl_table=11
/product="D-alanine-D-alanine ligase"
/protein_id="AAO07245.1"
/db_xref="GI:27358295"
/translation="MMRVNVLICGSGSSEHEISSLSEYVLOOQGLIENVNVLYKEIK
NMGNDQKRLVYLDIHRTKSVKSDSEFNHSIHDFVPTIHGFPEBTGDIQSLFELAI
PVLGQPEKSSFNKTSKLMYDAIGLPNPPYVFLTENNESHQLSREAFDKDKKXLF
VVAAGQSGVGCYSTKIEQLSDAIDAKAFGFSHQVLYEKAVKPRELVEVAYENNGQLH

gene	11080. .11454	/locus_tag="VV20281"	ISXGEETIAPDGAPEYSYDEKXSAGHSHTSEVAKNLTQEOLATIQLCSEVFPQNMRL HLSDRIPELTSEGETLYLNEVNTFPQMTIKI SMFPKMLQNHGHKFEHFLADCIERSL"	
CDS	11080. .11454	/locus_tag="VV20281"		
		/codon_start=1		
		/transl_table=11		
		/product="Conserved hypothetical protein"		
		/protein_id="AAO07246.1"		
		/db_xref="GI:27358296"		
		/translation="MAISHYGFNLQSPNDINIMOVKIKNRIQGLSLAVKSIDMW CDTASIIDPREFSSLDKQEARSGQAEFNFGALKNIDGERNAWICFPNGKIKSGSK AAIQHIEAVLIARAKAEQKK"		
gene	11457. .11768	/locus_tag="VV20282"		
CDS	11457. .11768	/locus_tag="VV20282"		
		/note="elf-1/SUT1; COG0023"		
		/codon_start=1		
		/transl_table=11		
		/product="translation initiation factor 1 related protein"		
		/protein_id="AAO07247.1"		
		/db_xref="GI:27358297"		
		/translation="MTLVYSTETGRIKKEEPPASRPKDGIVRIQETGRKRGKVCV VTGDLDPAPKLTAAELPKIKCGCGSVKDNIETIQGDARDKIKTHLEKKGMKPEAG G"		
gene	12360. .12782	/locus_tag="VV20283"		
CDS	12360. .12782	/locus_tag="VV20283"		
		/note="COG0499"		
		/codon_start=1		
		/transl_table=11		
		/product="S-adenosylhomocysteine hydrolase"		
		/protein_id="AAO07248.1"		
		/db_xref="GI:27358298"		
		/translation="WDFEAFNVLFWRVPPSIIKAGAEYFHVQPAYVRMLSGSIPINPM		
Query Match	39.9%;	Score 555.6;	DB 1;	Length 300883;
Best Local Similarity	64.9%;	Pred. No. 9.6e-97;		
Matches	876;	Conservative	0;	Mismatches 459;
			Indels	15;
			Gaps	3;
Qy	25	TCGAGAAATTTGTTATGACGTTATCGGCTTCGCTAGCCGTTGGGATTGACGATTGCG	84	
Db	118404	TCGATGAAAAAGATGACGATTGGCGGATTAACAAGTTGGTTGGCGTTGCTTACCG	118345	
Qy	85	CCCGATGCGGGGCACTTGGCAAGCTCGAAAAAGAAAGATTGAAATTCGGCTTCATCAAG	144	
Db	118344	GTTTGACTCGCAGAAATCGGCGCAGGCTGAAAAAGGACCTTAAAGTTTGGCTTATACAA	118285	
Qy	145	CTCACCGACATGAGCGCGCTGGCGGGTGGCGCGCGGAAAAAGGCTTCTTCAGACGAGGGC	204	
Db	118284	TTGACCGACATGACCGCGCTGGCGCATCGCTTATGAAAAAGGATTTTTTGGAGATGAAGGG	118225	
Qy	205	CTGTTCTGTCAACTGGAAGCGCAGGCCCAACTGAAAGTGATGATGATAGGATCGTGAAT	264	
Db	118224	CTGTATGTGAGCGCTTGAAGAGCGGACCAACTGGAAGTGCTGCTCGATCGCGATTGAT	118165	
Qy	265	GAGCACTGGAAGCGGCTCGCACATGCTGGCGCGCGCGCGCTTGAAGCGGCGGTTGCTTC	324	
Db	118164	GATGAACTGAGATGCGCGCCACATGCTGGGACGCAACCACTGGGTCCACATCGGCATT	118100	
Qy	325	GCGACCAAGCCGCAATTCGAGGTGCGCTTCACACATGAGGCTTCAACGCGCAAGCGGATTACG	384	
Db	118104	GSCACGCAAGCCCAAGTATATCCGCCCATTCAGATATGATCTCAATGATGAACGCGATACC	118045	
Qy	385	GTTGTCATGAATTTGGCGATCAGATGAAGCGGAACATACCGCTGGA---AGCGGTTAA	441	
Db	118044	GTTGTGAACGATGTTTGGCAGCAATGAAGCCCAATTTGGTCAAAAGCACTGACGGTTAA	117985	
Qy	442	CCGGTGATCCGATCAAGCAGATTATCTAAACCGGTGTGCAAAAATTCAAAGCCGAA	501	

Db	117984	CCCGTTACCCCATCAAAAGCCGATGCCCTTAAACCTGTGTAAACAAGTTTAAACCGGACACAA	117925
Qy	502	GGCAAGCGCTTCAATATATGGCGATGACCTTCCGGCCGGATCTCAACAACATCAACTGGCT	561
Db	117924	GGCAAAAGCCCTTCAATATATGGGACATGGTGTGTTCCCGTTTCAACGCATATACTAAGAACTGGCC	117865
Qy	562	TACTGGCTGGCGGCTGGCGGGTATCATCCCTGGCTACTATTTCCGCGCGCGA---GACATT	618
Db	117864	TACTGGCTTCGCGCTGGGGGATTCATCCCTGGCTTTTATATGCGCGCATTAAGCGGCAAC	117805
Qy	619	TCCGCGCAAAATCGCGCGACAGCGCTTGTGTGCGGTAGACCCCGCGCGCAAAATCCGCTCC	678
Db	117804	AGTGGCGAAATTTATATGTCGACGCTGCTCAAGGTACGCGCGCACACAAATGCCAGCC	117745
Qy	679	ACCGTGGAAAGCCGGCACCATTTTCGGTATATGCGTGGCGAGCGCTGGAAACGACAGGCC	738
Db	117744	ACCATGGAAAGCCGGGACCATCAAAAGGCTATTGCGTGTGGAAACCGTGGAAACGACAGGCC	117685
Qy	739	GTGTTCAAGGGGATATGGCGTGGCGGATCAACGATGAAGAACTCTGGAAGGACAGCGCG	798
Db	117684	GTGTTTAAAGGSGATTTGTGTGCCCCGTGTGACCGCATATATGATCTGGAATAAACACCA	117625
Qy	799	GAATAAGTCTTCGCGCGTGACCAAAACATATGGCGGAGAAATATCCCAACCTTATCTGGCG	858
Db	117624	GAGAAAGTATTTGGCGCTTGACAAAGTTGGGCGGAAAGTACCCCAATACCAATATCCGC	117565
Qy	859	GTGACCAAGCGCTGATTTCCGGCCCGCGATCTGGCTGAGCGCGCACATATACAAGAACCG	918
Db	117564	GTGTCAAAAGCTCTGATCCGTGGGGGCAATGGTGTATGATAAAGCATTAACCGCAATCCG	117505
Qy	919	AAGGAAGCATTGGAATGCTGGCGCAAAAACATACGTGGGTCTACGTGGAAGTGTCTG	978
Db	117504	GCCGAAGCGGTAAAGCTGCTTTCAACGACAGATACGTGGGCGCCGATCGGAAGTGAATC	117445
Qy	979	GCGGCTAGCATAGACGGCACCTTTCGAATACGAATAAGAGATTAACGGCGGCTAACCGGCAC	1038
Db	117444	GCCAAATTCATATACATGCGCACGTTTGAAATACGAATAAGCGATAGCGACAGTCCCGAT	117385
Qy	1039	TTCAACACCTTCTTTTCGCACACGGCGCAGCTTACCTCTTACAGCAGTCCAGTCTGTAT	1098
Db	117384	TTTATATGTTCTTTTCGTGCAAAACGCACTTATCCTTACTTACAGCGAGCGGATTTGTAT	117325
Qy	1099	CTGACCCACGCTGAGGGGCTGGGGCATGATCAATGAATTCAAAACCGGACAACTGTGATCTTG	1158
Db	117324	CTCACACAATATCGCCGTTGGGAGCAAAATCGAACATCAAGAAACCGGACCTCATGTGTCATGG	117265
Qy	1159	GATACCGGCANAAGGTCTACAGCCCGCGACATCTATCTCGCCGGCGCCAAAGAACCTGGTC	1218
Db	117264	GACATTCGCCACAGAGGTGTATGACCAGATTTTATCAACGCGCGCAGAGACATTAATC	117205
Qy	1219	GCAGAAGCCAAAGGCCAAGGCCAAGACCTTCCCTGCC-----GATACCTGCATCAAG	1269
Db	117204	GAAAGAAAGGACATATGAGCCGCACTGATTTTCTGACTTTGGCCAAAGAAAGCGGTTTTCGC	117145
Qy	1270	CCGTGCGAAGATTTCTTCATGACAAAGTGCCTTGCATGACCAACAGCCCAACGATTAC	1329
Db	117144	CCACCGCAAAACCCACCTTCACTGATACAGATCCGCTATGACGGCGCGTACCCCAATGTGAT	117085
Qy	1330	CTGCGCAAGTTTGGCATATAGCTCTGAAAGGC	1359
Db	117084	CTGACAGATTTGGCATATGGCTCTGAAAGGC	117055

RESULT 3					
AP005348/c	AP005348	248650 bp	DNA	linear	BCT 06-JUL-2004
LOCUS	Vibrio vulnificus y016	DNA, chromosome II, complete genome,			
DEFINITION	section 5/8.				
ACCESSION	AP005348	BA000038			
VERSION	AP005348.1	GI:37201116			
KEYWORDS					
SOURCE	Vibrio vulnificus y016				
ORGANISM	Vibrio vulnificus y016				

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Chen, C.Y., Wu, K.M., Chang, Y.C., Chang, C.H., Tsai, H.C., Liao, T.L., Liu, Y.M., Chen, H.J., Shen, A.B., Li, J.C., Su, T.L., Shao, C.P., Lee, C.T., Hor, L.I. and Tsai, S.F.	Comparative genome analysis of <i>Vibrio vulnificus</i> , a marine pathogen	Genome Res. 13 (12), 2577-2587 (2003)	2	(bases 1 to 248850)	Chen, C.Y., Wu, K.M. and Tsai, S.F.	Direct Submission	Submitted (29-MAY-2002) Shih Feng Tsai, National Health Research Institutes, Division of Molecular and Genomic Medicine; 128, Yen-Chiu-Yuan Road, Sec 2, Taipei 115, Taiwan (E-mail: petals@nhi.org.tw, Tel: 866-2-8146-1041, Fax: 866-2-2789-0484)
								This sequence was determined by the Sequencing Core of the National Yang-Ming University Genome Research Center (YMCG; http://genome.ym.edu.tw).
FEATURES	source	Location/Qualifiers						
		1. .248850						
		/organism="Vibrio vulnificus YJ016"						
		/mol_type="genomic DNA"						
		/strain="YJ016"						
		/db_xref="taxon:196600"						
		/chromosome="1"						
		383. .1786						
		/gene="VVA0914"						
		383. .1786						
		/gene="VVA0914"						
		383. .1786						
		/gene="VVA0914"						
		/codon_start=1						
		/transl_table=11						
		/product="anaerobic Ca-dicarboxylate transporter"						
		/protein_id="BAC96940.1"						
		/db_xref="GI:37201117"						
		/translation="MKRVYFPCITRENETITSLRINNMLYEPLFLIMLYIGSRYGGLGVSGSLGIVIEFVFKMPTSPPPVVMILIAVVCASILEAGGLKMTLQVAERMLRKRPKRVTIAFVTVTMTFMTGLAVSIMPDIIGDVALKNGIRERPMMAASVAQIILITASPISAAVVYLAQSDINHETLLSILVTPALFGLMLSLYSMRGKELEDDPEYQARIDKDPVMEKRLINTATSLDEHLPAARNSVLLFSLITIVIMIPERRITIGETKPIISMAVIVDMMLCTGGIILLATKTRDPSRVPGVFKSMAAIAIFGIAMSDITFYQVAMPQKSGIVEMVNNVPTFTALALIVSVVNSQAARMLPVGLGLALBMDLVGLMPALPYGFFIPNYPSDIATVFNMTGTTKIGKWFNHSFMSVGLIGVIGACCCGYLGGMIIG"						
		1858. .1891						
		/gene="rtpc2T08"						
		1858. .1891						
		/note="2 units of 17 bp in-tandem repeat"						
		/rpt_type=tandem						
		/rpt_unit="ttattctgcgtctaac"						
		complement (1905. .3185)						
		/gene="VVA0915"						
		complement (1905. .3185)						
		/gene="VVA0915"						
		/note="identified by Genemark and Glimmer2"						
		/codon_start=1						
		/transl_table=11						
		/product="conserved hypothetical protein"						
		/protein_id="BAC96941.1"						
		/db_xref="GI:37201118"						
		/translation="MMKRRILTLTLTQWLMPAVLAAGTNSQVYARVWMAQQALADNNLAALAIKLNAAEVSPPDVVALSKLGIYVQAQEPFRSLASLEKSLFELTFEBEAVTATKLTOLITQGDYKAIPLNKLIGYQBPENSAABEIRVYLAQAQSVENNRITYATGFORVREKPPPTTISLALAHVQLEOKSVYIPAKOILAEPSKQVMDAYNAVINTGQESALDMLTLBELQGLTNEKETSIALVIAAQGYQFQAQGLTARIESQOQDKLPKQAOVQAQKEWQAQALIEQRLAKESQYQWVALLENQLOLYQOVLISAEKMPNRQFPAQLAKTINALYRLERIDQAAVQAQKADALASTKOTQWMLKFLSHKLSAESTLDKAGSEVEREHLTGQRRVAVADKSDSS"						
		complement (3182. .3820)						
		/gene="VVA0916"						
		complement (3182. .3820)						
		/gene="VVA0916"						
		complement (3182. .3820)						
		/gene="VVA0916"						

```

/codon_start=1
/translate_table=11
/product="periplasmic protein TonB2"
/protein_id="BAC96942.1"
/db_xref="GI:37201119"
/translation="MRRLVAFPPVALALVAFSLAMALLVNSGKQALIDQTPSLFNM
VAVEQTEVTRRSQVPPPEIPEPQEPETPKMOTPMQVSTOATVETIDMRFNVSG
LAVSIPTVSAPTONASNTTKOLMPLSRVAVYPAKAKRRIGEVYLKFDIDENGR
KRIEVEAQPAPPEENMDAIKRMRYQPIYDGAQITTYGTTKIEBFMAQ"
complement(3820..4224)
/gene="VVA0917"
complement(3820..4224)
/gene="VVA0917"
/codon_start=1
/translate_table=11
/product="biopolymer transport protein ExbD2"
/protein_id="BAC96943.1"
/db_xref="GI:37201120"
/translation="MRIGRQSGSDEAQIDLTSMLDIVEFMILFPIVTSFVRESGVE
VNRPOANYSAQKDAGI FVAITAANDIYIDKRMVDERVQASIERLLLEQPDASLVIO
ADEYANGVYVYMDAKAGVNTALAEKS"
complement(4221..4787)
/gene="VVA0918"
complement(4221..4787)
/gene="VVA0918"
/codon_start=1
/translate_table=11
/product="biopolymer transport protein ExbD2"
/protein_id="BAC96944.1"
/db_xref="GI:37201121"
/translation="MIMDLHSDIMPBELTSSWMNALYFMOQGEVLMLAAVAVI
FMLTEKLFIYCAVPRRAETLOKMRHDLISWFAHQREANLIDHLOLFHILA
FKIVAVICPMIGLGTVMGISVDFVMAVTAQSSNPKMAAGISLATLPTMAGVAAAL
VGMFYHARIVYKCONREYVLEOSLRSDS"
complement(4784..6151)
/gene="VVA0919"
complement(4784..6151)
/gene="VVA0919"
/codon_start=1
/translate_table=11
/product="biopolymer transport protein ExbD-related
protein"
/protein_id="BAC96945.1"
/db_xref="GI:37201122"
/translation="MNKCRLLSIALIVSLSPFAGTOLVPRKATQESROOAEONORE
NAFLSQEQDLAKRQLEKQSTLBEKNALISQPSDNEVTYSHLBEELRYTGSIGE
LPGVVRQHAKELEKADINDAVTAIGQTEQVETVNOQVAAQSLPDIAQLKTLHWQOOI
VASGEMRALQVNLINAEQAVEPTDVNALGFALNQGFPNKSDFGFIPIPAKLPEG
APTLDLRSQAHLSIMVPLVIDPTRGVILAEQVASHPSLADRFAGAGVGKVFALALVG
LVIALYRGILLTTKSKIRLOLKOPNQNPPLGRVLAVYDKQSPRSVALRLILET
IYDEQTEKEKGLSMCLKLAALPMLGLGTVMIGLETFQVITQFQVITQFGNDPKWASGISM
ALVTYVGLGVAAAMPILAHNVLSGMAEIRSLBEKQIGLVAQAEQQLPSLPTMDK
VGNAA"
complement(6152..6919)
/gene="VVA0920"
complement(6152..6919)
/gene="VVA0920"
/feature="identified by Genemark and Glimmer2"
/codon_start=1
/translate_table=11
/product="conserved hypothetical protein"
/protein_id="BAC96946.1"
/db_xref="GI:37201123"
/translation="MSGYKAGAAALLIMVSPGAMSNDEQAQKIQNRRTTASQSV
INASADQLOAEIEILKEBYRNIELYRHLISMANOEKSSDAEKIQDIQVTRQ
GIVPLMYDLTKTTINQDMPIKWQQRERIETKEQNMVADVADAEKFRILQYTRQ
IELDYSKLSAYOAPITTLADGQIEVEQHLGRVSLIATLLANQFMYVDQQRHMLS
LTTIPSTDVNRASFDAKQVAPAMLMVPALATAEAK"
complement(7182..7670)
/gene="VVA0921"
complement(7182..7670)
/gene="VVA0921"
/codon_start=1

```

```

/translate_table=11
/product="uncharacterized conserved protein"
/protein_id="BAC96947.1"
/db_xref="GI:37201124"
/translation="MRVMTSRVAFVVDVQNIYYTCCKERYRHFYDNRFWQATQDKQV
VVAHAYVASSDPCGRQGFHILIRGIGFVVKKPFQORDGSAKGMVDGLADAEHA
PMVEVILLSDGDFEILVERIQGFQFKVTVYGVPLTAQNLQAVVDKYPIDELL
LP"
7796..8227
/gene="VVA0922"
7796..8227
/gene="VVA0922"
/feature="identified by Genemark and Glimmer2"
/codon_start=1
/translate_table=11
/product="hypothetical protein"
/protein_id="BAC96948.1"
/db_xref="GI:37201125"
/translation="MRYSIVILIFSLAHSVPADSEKMRHSIMFAPSHDOYVQOF
LLETMECSLOERDITVLTKEDGSTVPMVAQDAFDLSVLAQVFDIPKQHTGLVG
KDGRKLEMGKRTMTLLKVTIDQMLPQOEMQRNSRCAT"
complement(8234..9076)
/gene="VVA0924"
complement(8234..9076)
/gene="VVA0924"
/codon_start=1
/translate_table=11
/product="uroporphyrinogen-III methylase"
/protein_id="BAC96949.1"
/db_xref="GI:37201126"
/translation="MLIRSKHFPREALSGTLVHQIYISGTMTLTSLSLQNOGFS
LVGAGGPDLLTVGYVIOQAEVVVDRLVSPRIILANAEKAMIYVKKLIDFHCV
POEINQILVRALEKGRVRLKGGDSITPFRSGEELBELHGVRFVDPITAAQ
ATAVAGIPLTRHDAQSVQPTTGHQKXGKLEMPSLQANNTLVFVGLKOCRYIAE
KLENGLDASTSCALIENTRBOQVRLTSLQPELTAQALISPLIYVGSYTSLHDK
LKW"
8255..8395
/gene="VVA0923"
8255..8395
/gene="VVA0923"
/feature="identified by Glimmer2"
/codon_start=1
/translate_table=11
/product="hypothetical protein"
/protein_id="BAC96950.1"
/db_xref="GI:37201127"

```

Query Match 38.9%; Score 541.2; DB 1; Length 248850;
Best Local Similarity 64.2%; Pred. No. 5,7e-94;
Matches 867; Conservative 0; Mismatches 468; Indels 15; Gaps 3;

QY	25	TCGAAGAAATGTTATTCGATTCGCTTCGCTTACCGCTTGGGGTTTGACGATGGC	84
DB	46865	TCGATGAAAAGATGATGACAGTTGGCGTATTAACAAGTTGGTGGCTTCACCGC	46806
QY	85	CCCGATGTCGGGGCGTTGGCAAGCTGGAAAAAGAAATTGAATTCGGCTTCAACG	144
DB	46805	GTTTACTTCGCAAGAGTCGGCGAAGCGCTGAAAAAGAGACTTAAGTTTGGCTTACAA	46746
QY	145	CTCAGCGACATGCGCGCGCTGGCGGTGCGCCGCGAAAAAGCTTCTTGAGAGACAGGCG	204
DB	46745	TTGACCGACATGCGCACCGTTGGCGATTCGTTGAAAAAGATTTTGGAGATGAAGGG	46686
QY	205	CTGTCGTCGCAACTGGAAGCCGAGCCCACTGGAAGTGTGATGATGATGAGGTTCGAT	264
DB	46685	CTGTATGTGACGCTTGAAGCGGAGCCCACTGGAAGTGTGCTGATCGCTGATGATGAT	46626
QY	265	GGCGAAGTGAAGCGGTCCCACTGCTGGCGCGCGCGCTTACGCGCCAGCGTTGGCTTC	324
DB	46625	GGTGAACGTGATGTGTGCGCACATGCTGGCAGAGCCCAACACTGGGTGTGCCACATCGGCATT	46566
QY	325	GGCAGCAAGCGCGATATGAGAGGTGCGTTACAGATGGCGCTTCAACGCAACCGATTACG	384

Db	46565	GGCAGCGAAGCCAAAGTATCAACCGCATTGATGTGATCTCAATGTTAAGGATACACC	46500
Qy	385	GTGTCCATGAAATCTGGCATCTAGTGAAGCCGAACATACCGCTGGAAAG--CGTTAAA	441
Db	46505	GTGTGAAACGAAGTTTGGCAGCAATATGAAGCCCAATTTGGCCAAAGACATGACGGTTAAA	46446
Qy	442	CCGGTGCATCCGATCAAGGCAAGATTATCTTAACCGGTGCTCCAAAATTCAAGCCGA	501
Db	46445	CCCGTTACCCCATTAAGACCGATGCCCTTAAGCCCTGTCTGTAACAGTTATTCGCGACAA	46386
Qy	502	GGCAAGCCGTTCAATATATGGCGATGACCTTCCCGGCCGATCTCAACAATCAATCGCGT	561
Db	46385	GGCAAAAGGTTCAATATATGGGATGAGTGTTCCTCCGTTTCAACGATATCTACGAATCGCC	46328
Qy	562	TACTGGCTGGCGGCTGGCGGTTATCAATCTGGCTACTATTTCGCGCGCGA--GAATTT	618
Db	46325	TACTGGCTGGCGGCTGGCGGAGATTATCTCTGGCTATTATGCGCGGCTAAAGGTGCAAC	46266
Qy	619	TCCGGCCAAATTCGGGCGAAGCCTTGTGTGTGCGTGAACCCGCGCGCAATATGCGCTCC	678
Db	46265	AGTGGCCAAATTAATGCGAGTGTCTCTACGCTCAAGCTCTACACCAAAAGGCCAGCC	46206
Qy	679	ACGCTGGAAGCCGCGACCATTTTGGTTATTGCTCGGCGAGCCGTAACCAAGCAGGCC	738
Db	46205	ACCATGGAAGCGGGGATACCATTAAGGCTATTTCGCTGTGTATACCTTGAACCAACAGCG	46146
Qy	729	GTGTTCAAGGGCATAGCGGTGCGCGGTATCAACGATGAAGATCTTGGAAGACAGCGCG	798
Db	46145	GTGTTAAGGGTATTGGGTGCGCTCGGTGTGACGCACTAGATCTGAAAAAACACCA	46086
Qy	729	GAAAAAGCTTTCGGCGTGAACCAACAAATGGGCGGAGAATATCCCAACCTATCTGGCG	858
Db	46085	GAGAAAGTATTGTGTGTGACAAAGATTGGGCGGAAAAAGTACCCCAATCCCAATATCCCG	46026
Qy	859	GTGACCAAGCGCTGATTTCCGGCGCGCATCTGGCTGGAAGCGCGCAATTAACGAAGCCG	918
Db	46025	GTGTGCMAAGCCCTGATCCGTCGGCGGCATTGGTTAGATGAACATTAACCGCAATCGC	45966
Qy	919	AAGGAAGCCATCGAATATGCTGGCGCAAAAACAATACGCTGCTGCTGACGTGAAGTCTG	978
Db	45965	ACCGAAGCCGTGAAGCTACTTTCAACGAGTAGTAGTGGCGCGCATGACGAAGTATC	45906
Qy	979	GGCGCTAGCATGAACCGCATTTTGGAATACGAAAAAGACGATTAACGCGCGCTACCGGAC	1038
Db	45905	GCCAAATTCATGACTGCGACAGTTTGAATACGAAAAAGGCGATTAAGCGAAGATGCCGAT	45846
Qy	1039	TTCAACAACCTTCTTGGCGACGGGCGCAGCTATCCGCTTACAGCAGTGCATCTGTGAT	1098
Db	45845	TTTAATGTGTCTTTCCTTCAACAGCCACTATTCCTACTACAGCAGCCGATTTGGTAT	45786
Qy	1099	CTGACCAAGCTGAGGCGCTGGGCGCATGATCAATGAATTTCAACCGGACCACTGTATCTG	1158
Db	45785	CTCACAACAATATGCGCCGTTTGGGACAAATGGAACACCAAGAACCGGACTCATGTATCATG	45726
Qy	1159	GATACCGCGAAGAAAGTCTACCGCGCGGACATTTATCTGCGCGCGGCGCAAGAACTGTGTC	1218
Db	45725	GACATCCCGAAGCAGGTGTATTCGACACGATATTATTAACAAGCGCGCAGAAAGCGCTAATC	45666
Qy	1219	GCAAAAGGCAAGCGCCAGCGCGGAAGACTTCCCTGCG-----GATACCTGCATCAAG	1269
Db	45665	GAAATAAGGACATTTGAAGCGCGCAGAGATTTCCTGACTTTTGGCAAAAGAACCGGTTTTGCG	45606
Qy	1270	CCGTGCGAAGATTTCTTCATCGACAAAGTGCCTGTGCATGCCAACAAAGCCCAAGCATTAAC	1329
Db	45605	CCACCGGAACCCACTTCAATTGATACAGATCACTATGATAGCGCGCGTAAGCCCAATGCTAT	45546
Qy	1330	CTGCGCAAGTTTGGCATAGTCTTGAAGGC	1359
Db	45545	CTGCAGCAGTTGCGCATTTGGCTGAAAGGC	45516

LOCUS	AP005087	302100 bp	DNA	linear	BCT 07-APR-2004
DEFINITION	Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 4/6.				
ACCESSION	AP005087	BA000032			
VERSION	AP005087.1	GI:28809161			
KEYWORDS					
SOURCE	Vibrio parahaemolyticus				
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.				
REFERENCE	1 Naeu, H., Iida, T., Sugahara, T., Yamaiichi, Y., Park, K.S., Yokoyama, K., Makino, K., Shingawa, H. and Honda, T. A filamentous phage associated with recent pandemic Vibrio parahaemolyticus O3:K6 strains J. Clin. Microbiol. 38 (6), 2156-2161 (2000)				
JOURNAL	20295086				
MEDLINE	10834969				
PUBMED					
REFERENCE	2 Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T., Tagomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A., Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shingawa, H., Hattori, M. and Iida, T. Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V cholerae Lancet 361 (9359), 743-749 (2003)				
TITLE	3 (bases 1 to 302100) Hattori, M., Yamashita, A., Oshima, K. and Shiba, T. Direct Submission Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@genome.is.kitasato-u.ac.jp, URL:http://genome.is.kitasato-u.ac.jp/ Fax:81-42-778-8193)				
COMMENT	This work was done in collaboration with Kozo Makino, Ken Kurokawa, Katsushi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima, Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga, Takeshi Honda, Hideo Shingawa, Tetsuya Iida (Osaka University), Yoshio Iijima (Kobe Institute of Health), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science. This clone was isolated from a patient presenting with acute gastroenteritis.				
FEATURES	Location/Qualifiers				
SOURCE	1. 302100 /organism="Vibrio parahaemolyticus" /mol_type="genomic DNA" /strain="O3:K6" /sub_strain="RIMD 2210633" /db_xref="taxon:670" /chromosome="2" 53..1139 /note="disrupted by frameshift, similar to part of hypothetical protein GB:AAF96452.1 (AE004386) percent identity 62 in 165 aa" complement(1224..2012) /gene="VPA0861" complement(1224..2012) /gene="VPA0861" /note="similar to GB:AAF96483.1 (AE004389) percent identity 73 in 261 aa" /codon_start=1 /evidence=not_experimental /transl_table=11 /product="conserved hypothetical protein" /protein_id="BA062204.1" /db_xref="GI:28809162" /translation="MKVWIKASALLTLAGTATCASAPGGRNQTLMFSDQMSLAKS PDMKRDIPISDKDTNAYVCCVAKOITNTLPQGFKQMEVAVFSDVNAFALPGG KIGVYGLKLVAKNDQLTAVIGHEVAHVAIAHSNERLSQSLANVAGSLANVAIGAS EKVQGLQALGVGVGVVILPYGRQDEADIVGLVEYAKAGFNPQSVDLNKKM SASGCAQCPPEFSTPSTHSRTIKDLOATINTLPCKNVAPKCG"				
misc_feature					
gene					
CDS					

```
gene      2217. .2651
CDS       /gene="VPA0862"
          /note="similar to GB:AAF96482.1 (AE004389) percent
          identity 75 in 144 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="BAC62205.1"
          /db_xref="GI:28809163"
          /translation="MSIWKKPISLDILNATSRNTWMEHLQIEYTFDLSATMPVC
          SFTHOPLGMHGASVLAETLGSIANFCSEGVNCGLDINAHVRAMRGVYIGT
          AKPIHGVSTOWQIETIBDRRLVCTSLTIAVQOKKAKLA"
          2594. .2927
          /gene="VPA0863"
          /gene=".2927
          /note="similar to GB:AAF96481.1 (AE004389) percent
          identity 56 in 76 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAC62206.1"
          /db_xref="GI:28809164"
          /translation="WYIERSLGKIIATQREIYIKLTGEMVTLQAQTDALQLLGKGN
          VVLAHGACKKSIKLDNEQLTQLAOFIIGIDIQ"
          3007. .4713
          /gene="VPA0864"
          /gene=".4713
          /note="similar to GB:AAF96480.1 (AE004388) percent
          identity 66 in 565 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="putative ABC transporter periplasmic solute
          binding protein"
          /protein_id="BAC62207.1"
          /db_xref="GI:28809165"
          /translation="MSSPLRLVQFETLPEKFAQNDTEVQLIEDITDALFCTRRNARIVL
          NKLSEGMIEWHPAAGRGKLSLIFKGRSDVSENLAARDYDEGKIGQLDALNDAA
          RLTOVIGYLGLOHROGEQVRLPYRPLAMLNPOKPMRBSQHTAROVPSGLTRLE
          SEOLOPLAHTEALSDTHMRPYLRGCVFHNGBELTTTCVVESTIALSGNLFSHIQ
          QVSSPQPTVDIKLVRPDRYLPLALSESQAKVLPQALSEDFDRKPIGTGFPQYKVN
          DKRLILNATGIFADPRVAEFLSQTLNHHQYAHVPQKVLGLGLQAFGLKPMID
          LRPTSGEVPOQKSTIVAYQKKHMPFVIAKAIKHLKPHSVDFEITDYQPPNP
          EDVWVKAMGIATNRDIALGMLLDVSDIEKFSAGYDFSEMATLVDMQWAGIHADPFA
          RELGRQLVKSCVYIPHFHGMGLVGNKHSAQLNAKCNALGMWDFSNVWVKPEIEDNGQ
          AE"
          complement (5251. .5847)
          /gene="VPA0865"
          /complement (5251. .5847)
          /gene="VPA0865"
          /note="similar to GB:AAF96687.1 (AE004407) percent
          identity 71 in 192 aa"
          /codon_start=1
          /evidence=not experimental
          /transl_table=11
          /product="conserved hypothetical protein"
          /protein_id="BAC62208.1"
          /db_xref="GI:28809166"
          /translation="MKSTSTLLGLTIFVAIVFLFSALFESSRAVWFAEIIPIAGILIA
          IMAISLYQPSNTAYLIMFIMLCLHTIGAKYFAEVPFMPNNLLGSENNNDRAVAF
          AGLYAYPIAEYLIRNKKENPTFSCWFAIPALMSLAAGYEIIEMWYAEIAGDEGIAF
          LSGQGDIMDAQMLCDTTGAILSLFLMSAQRRFAKPF"
          6123. .6962
          /gene="VPA0866"
          /gene=".6962
          /gene="VPA0866"
```

```
/note="similar to GB:AAK22187.1 (AE005694) percent
identity 30 in 298 aa"
/transl_table=1
/evidence=not experimental
/product="putative oxidoreductase iron/acorbate family"
/protein_id="BAC62209.1"
/db_xref="GI:28809167"
/translation="MKLETVDYIADDAQAQFVASLRENGFGLKXNHPIDBELVESIYK
NMHFERCEERKNDQFQNVETODGYKPPPSVSAKCHKYKDKIEFYHYPKQIPROK
ABILEYTERANAFROELIGWVEYAPKDVQEKFSIALSEMINNSDKTLIRLYAHYPT
GSEEPGAIRAAHEDINLITVLPANEEGLQVMSKDGSEIDVPCDFGULINIQMLO
BASGGFSPSTHRYINIPGARQEKSRISLPLFHPKPTVLSERYTAHSYLMERLEL
GVI"
7095. .7715
/gene="VPA0867"
/gene=".7715
/note="similar to GB:AAF96686.1 (AE004407) percent
identity 63 in 204 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="DnaU-related protein"
/protein_id="BAC62210.1"
/db_xref="GI:28809168"
/translation="MEDQYGVCDSSQSEIENPLMPLEILIKROPSGMKVHTLASKLS
ERGYNOIDPSPDQDLFRNPLINMALYQLOETLPPDSWLQVEAMDILMSALEASRN
SYDENDPLRQYTLDMSHYEANEDERYRLINFEWTRVQCFVGASASHNIDKAKALFLD
LAGDVSRTERRQWRKLLRNWHPDENGDARFVYLCEAMVYLKNG"
complement (7844. .9499)
/gene="VPA0868"
/gene="(7844. .9499)
/complement (7844. .9499)
/note="similar to GB:AAF96684.1 (AE004407) percent
identity 78 in 548 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC62211.1"
/db_xref="GI:28809169"
/translation="MDQLTATLKIEKQNYRAYOQIKQYDFTFLEFIDHYQSDPYA
SASRRAPRANSLGLEWLKDESPAFQAAEDFIARSEQPAKQNTVSIANGOTVL
DSTAYLPTBEGIELRFRVNLPAEGSVYLGKANNILTTHLPETIRATLEBIDKEAM
VKQOYVDSQALREQLAEHNLVAFVANGSVLPARGCDLPMKEAVFETPESIOVT
LHAPKGYVTGLGIPKGTLLVGGFHEKSTLINAIESYDHIIPGDREIYVYTOKA
MKIRAEADRQVHNHLSYINHLPEKQOTADFTQDASGTSQAAMLDESIAGATSL
LIDETSAITNMHIRBERMOALVAKDEBITPLVRIQGLRBLDISTIIVMGSGDYL
DVANTYIQMHQYQAVDYTEKAKOVIAQHTORHNSESLQTFRRALNRVALNNILT
DKPFRVSAKGDLSRFGKEPTDLSALBQIESADEVNAIGMLWFOALQLPWCNNPAKE
IEEMISGSMHNSLPKQGDIAKPRITLDVVAALNRNKKSGFKXSH"
complement (9571. .11481)
/gene="VPA0869"
/gene="(9571. .11481)

Query Match      37.1%; Score 516.2; DB 1; Length 302100;
Best Local Similarity 64.4%; Pred. No.3,6e-89;
Matches 827; Conservative 0; Mismatches 443; Indels 15; Gaps 3;

QY      101 TTGGCAAGCTGGAAGAAGATTGGAATTGGCTTCATCAAGCTCACCGCATGGCGC 160
          |||||
DB      217725 TGGGTGAACACGAATTTGAAGACCTGAAGTTGGTTATTATTAAGTTAACCGATATGGCCG 217666
          |||||

QY      161 CGCTGGCGGTGGCGCCCGCAAAAAGGCTCTTCGAGAGAGAGGCGCTGTTTCGCAACTGG 220
          |||||
DB      217665 CGTTAGCGGTGGCTTACAGAAAAGGCTTTTGAAGAAGAGGGCTATACGTTACGCTCG 217606
          |||||

QY      221 AAGCGCAGGCCAACCTGGAAGGTGATGATAGAGGTGCTGATGAGCGCACTGACGAGCT 280
          |||||
DB      217605 AAGCGCAGCAAACTGGAAGATGATCTTTGAGACCGCGTTATGATGATGTAACCTCGATGGCG 217546
          |||||

QY      281 CGCACATCTGGCGCGCGCGCTTATAGCGGCGAGCGTTGGCTTGGCCACCAAGCCCGAT 340
          |||||
```


Db 217545 CGCATATGCTGGCAGGGCAGCCATTTGGGGCGCAACGATTTGGTGTGGGACAAAGAGG 217486
 QY 341 TCGAGGTGCGCTTACAGATGGCTTCAACGGCAACCGGATTACGGTGTCCATGAAATCT 400
 Db 217485 TGATTAACGGTTCAGTATGTGATTTGAAACGGCAACCGAATACCTGTTTCCAAAGACACGT 217426
 QY 401 GGCATCAGATGAAGCCGAAATAC---GCTGGAAGCGGTAAACCGGTGCATCCGATCA 457
 Db 217425 GGGAAACGATGAAGCCACATACCAAGACCGCAGATGGCAACCTGTACATCAATCA 217366
 QY 458 AGCAGATTAATCTAAACCGGTGTGCAAAATCAAGCCGAAGCAAGCCGTTCATTA 517
 Db 217365 AAGCGGATTCGTGAAGCCCGTGTGACGTTATGCGCAAGAGGCAAGTATTCAATA 217306
 QY 518 TGGCATTGACCTTCCCGCGCGGATTCACAACTCAAACTGGCTTACTGGCTGGCGGCTG 577
 Db 217305 TGGGCAATGTGTTCCTGTTTCTACTCACAATTAACGATGTCGTTATTTGGCTTGTGCTG 217246
 QY 578 GCGGTATCAATCTGGCTACTATTTCGCG---GCCGAGGACATTTCCGGCCAAATGGGG 634
 Db 217245 GTGGGATTCATCCAGGTTTATTAACACAGACTCAGCGGATTAACAGTGTCAATTTGATG 217186
 QY 635 CAGACGCTTGTGTGTGAGTACCCCGCGCGCAATGCGGTTCACAGCTGGAAGCCGGA 694
 Db 217185 CTGACGTTCTTTTGAAGTTTACACCGCACCAAAATGCCAGCCAAATGGAAGCGGCA 217126
 QY 695 CCATTTTCGTTATTTGCTGTGCGGAGCGGTGGAACGACAGCGGTTCAGAGGCGATAG 754
 Db 217125 CAATCAAGGGTACTCGTGTGCGGAACATGGAACGAGCAAGCGTATTCAAAGGCAATG 217066
 QY 755 GCGTCCGCGGATCAACGATGAAGAACTCTGGAAGGACACGCGGAAAAAGTTCTGGGG 814
 Db 217065 GTACGCCATCTGTGACCGATTCGAAATCTGGAAGAACCAACGAGAAAGTGTGGTG 217006
 QY 815 TGAACCAACATGGGCGGAGAAATACCCCAACACTATCTGGCGGTGACCAAGCGCTGA 874
 Db 217005 TGTCCAAACATGGGCGGAGAAAGAACCCCAACGCAATTCGTGTCTGGAAGCCCTGA 216946
 QY 875 TTGCGGCGCGGATCTGGCTGAGCGCGACATTAACAAGACGCAAGAGCCATCGAAA 934
 Db 216945 TTGCGAGCGGACATTTGGTGTGATGAAGAACAGATGCCAACCGCAAGAGCGGTGAGA 216886
 QY 935 TGTGCGCGCAAAACAAATAGTGTGCTGAGTGTGAGTGTGCGGCTGAGATGAAG 994
 Db 216885 TTTTGGCGAAGACGAGTATGTGGAGCGGATGCCGATGTATTCGCAACTCAATGACAG 216826
 QY 995 GCACTTTCGAATACGAAGAAAGAGATTAAGCGCGCTACCGGACTTCAACACTTCTTC 1054
 Db 216825 GCACGTTTAAATACGAAGAAAGGAGCAAAAGAGACGTACCAAGCTTCAACGTTCTTC 216766
 QY 1055 GCCACGCGCGCGAGCTATTCGCTCTACAGAGTGTGATCTGTGACCCAGCTGAGGC 1114
 Db 216765 GTTCATTAACCCACGATCCCGTACTACAGGATGCGATTTGTGATCGAACCCAAATCGCC 216706
 QY 1115 GCTGGGGGATGATCAATGAATCAACCGGACAACTGTATCTGATACCCGCCAAGAAG 1174
 Db 216705 GCTGGGAGCAAAATCTCAAGCGAGAAATCCGATGATGTGATCAATGTGGCGAAGAG 216646
 QY 1175 TCTACCGCGCGGACATCTATCTCGCGCGCGCAAGGACTGTGCGAGAGGCAAGGCA 1234
 Db 216645 TTTACCGTCAATATCTACCAAGACAGCTCAAGCGTTAATGAAAGATGTGTCTTA 216586
 QY 1235 AGCGCGAGACTTCCC-----TGCAGATACCTCGATCAAGCGGTGCGAGAAATTTCT 1285
 Db 216585 GCAAAAGAACTTCCAGATTTCAAGTGTGAGATGATGTTTGTCTCGCCACAACACACT 216526
 QY 1286 TCATTCAGCAAGAGCGCTTCGATGCAAGAGCCCAAGATTAACCTCGCAAGTTTGCA 1345
 Db 216525 TTAATGACAACTCTGTTATGACGCTGCGAGCGCAATTAATATCTGGAGAAAGTTCTTA 216466
 QY 1346 TAGGTCTGAAGGCAAGCAAGCCGT 1370
 Db 216465 TCGGGCTTAAGGAAAAAGCAAAAGT 216441

RESULT 5
 BX294133
 LOCUS
 Pirellula sp. strain 1 complete genome; segment 1/24.
 DEFINITION
 BX294133 BX119912
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rhodopirellula baltica SH 1
 Bacteria; Planctomycetes; Planctomycetaceae; Rhodopirellula.
 REFERENCE
 1 (bases 1 to 311050)
 Gloeckner, F.O., Kube, M., Bauer, M., Teeling, H., Lombardot, T.,
 Ludwig, W., Gade, D., Beck, A., Borzym, K., Heilmann, K., Rabus, R.,
 Schlesner, H., Amann, R. and Reinhardt, R.
 Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1
 Proc. Natl. Acad. Sci. U.S.A. 100 (14), 8298-8303 (2003)
 MEDLINE
 PUBMED
 22735913
 12835416
 2 (bases 1 to 311050)
 Kube, M., Borzym, K., Heilmann, K., Klages, S., Marguardt, I.,
 Lehnack, S., Beck, A., Pawlik, R., Reinhardt, R., Gloeckner, F.O.,
 Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Rabus, R.,
 Schlesner, H. and Amann, R.
 Direct Submision
 Submitted (21-JAN-2003) Max Planck Institute for Molecular
 Genetics, Proscience Imneustrasse 73, D-14195 Berlin, Germany
 Max Planck Institute for Marine Microbiology
 Planck Institute for Marine Microbiology
 Bremen, Germany
 This project was carried out by
 *Max Planck Institute for Molecular Genetics, Berlin, Germany; *Max
 Planck Institute for Marine Microbiology, Bremen, Germany; in the
 framework of the REGX-project, <http://www.regx.de>
 Genome Center
 Center: Max Planck Institute for Molecular Genetics
 Center code: MPIMG
 ----- Summary Statistics
 Sequencing vector: pUC19; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 7142841 bases at least Q40
 Consensus quality: 7145138 bases at least Q30
 Consensus quality: 7145484 bases at least Q20
 Quality coverage: 8.03

 This sequence was finished as follows unless otherwise noted: all
 regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid sequence; assembly was additionally confirmed by long
 range PCR and cosmid end sequences.

 See <http://www.micro-genomes.mpg.de/pirellula/> for more information
 including minimal tiling path from a set of 220 cosmids out of
 908. See the misc feature tag below for the boundaries of the MIP
 cosmids. ----- Annotation
 Center: Max Planck Institute for Marine Microbiology
 Celsiusstrasse 1, D-28359 Bremen, Germany.
 Center Code: MPIM
 Email: fog@mpi-bremen.de
 Phone: +49 (0)421 2028 938 Fax: +49 (0)421 2028 580

 Three different programs (Glimmer, Critica and Orpheus) were used
 for ORF-prediction. A nonredundant list of ORFs was generated by
 suitable parsing of the results.
 Automated annotation was done with the software package Pedant Pro
 (<http://www.biomax.de>). All ORF predictions and annotations were
 manually corrected by considering all results of the different
 tools applied. See <http://www.regx.de> for more information and

access to supplementary information.

FEATURES	Location/Qualifiers
source	1..311050 /organism="Rhodospirillum rubrum" /mol_type="genomic DNA" /strain="1" /db_xref="taxon:243090" complement(3..182) /locus_tag="RB2" complement(3..182) /locus_tag="RB2" /locus_start=1 /locus_end=11 /product="hypothetical protein-signal peptide and transmembrane prediction" /protein_id="CAD71330.1" /db_xref="GI:32442812" /translation="MSRLFFSALFAFSVSFTVGGCGGNEVIEDTRSPAEVQOEOD YEKMEETWDSATNS" 39..260 /locus_tag="RB3" 39..260 /locus_start=1 /locus_end=11 /product="hypothetical protein-signal peptide and transmembrane prediction" /protein_id="CAD71331.1" /db_xref="GI:32442813" /translation="MLFVIFLLHFGWRAVGFDFVAAASATNSEDRNAERKQGEK QSRHYCSMDARNECEKLVRELLILRLISP" complement(399..1514) /locus_tag="RB5" complement(399..1514) /locus_start=1 /locus_end=11 /product="hypothetical protein-signal peptide and transmembrane prediction" /protein_id="CAD71332.1" /db_xref="GI:32442814" /translation="MTKRLPTRQRLVLAAGLITPLAASTAKAGFQWEMETGSPY YEDDAVYDSEWFDGNDYPTDEAIGRWNETYDDEALTSDDNDIDMTWKDGY ADTSSDITFYDYDHGYSDYGDNNNDGNYEYDNDGVDAFASVYDTQGD IYDDATYISFNSAEBNDOKTAKKADONASAKSOFSTITRAKKVPTATTNV VELQNSBEKSMIADGPADQIDRMPLNDITTVGSGPRKADTVLSRSIEHNGQ MVERSGREYSGTISTKTVDYRGQQRQAKIKTDGKMLVDLGRKODLTNITEG QYTINGPAVQVDRVWLATKVTLDGDTIEISRMQK" 1548..1706 /locus_tag="RB7" 1548..1706 /locus_start=1 /locus_end=11 /product="hypothetical protein" /protein_id="CAD71333.1" /db_xref="GI:32442815" /translation="MRSRSQSVQETNGKRVAKKDIGAARVVAQPNRFGKLSVAG RGGKQNVQ" complement(1733..5263) /locus_tag="RB8" complement(1733..5263) /locus_start=1 /locus_end=11 /function="DNA synthesis and replication; transcription" /note="Best DB hits: BLAST: p1r:D83545; probable helicase PA0799 [imported] - Pseudomonas; E=7e-95 emb1:CA67095.1; (X98455) SNF [Bacillus cereus]; E=8e-94 p1r:C72027; swissf family helicase 2 - Chlamydomonas pneumoniae; E=4e-87 COG: PA0799; COG0553 Superfamily II DNA/RNA helicases; SNF2 family; E=7e-96 PPM: PF02803; Thiolase, C-terminal domain; E=0.11 PF00176; SNF2 and other N-terminal domain; E=3.9e-64 PF00271; Helicase conserved C-terminal domain;

Qy	79	ATTGGCCCGATGTCGGGCGAGTTCGCAAGCTGGAAGAAGAAATTGGAATTCGGCTTC	138
Db	184764	AATCCACTGAGTCTGCTGCATGTTGACCTTGGAATTCGATTTGATTCATTCGGCTTC	184823
Qy	139	ATCAAGCTCACCAGCATGCGCCCGCTGCGCGTGCAGCCGCAAAAAGCTCTTCAGAGAC	198
Db	184824	ATCAAACTGACCGATCGGCTCTTGTGTAATCGCAAGAAGAGAGGCTACTTCAGACAC	184883
Qy	199	GAGGGCTGTTGTCGCAACTGGAAAGCGCAAGCTGGAAGCTGTAATGAGTTCGCTC	258
Query Match 31.0%; Score 431; DB 1; Length 311050;			
Best Local Similarity 62.6%; Pred. No. 8.9e-73;			
Matches 747; Conservative 0; Mismatches 425; Indels 21; Gaps 4;			
E=1.2e-24			
/codon_start=1			
/transl_table=11			
/product="helicase"			
/protein_id="CAD71334.1"			
/db_xref="GI:32442816"			
/translation="MDWIAEDQSLGILMGSAVERLVSDHQFDSRAAMLTVKPT LTQDNGVAFDFQVWNEKQOTVALIEVTPQVEDDPGAGELSECCGCVAFPLRDN SCPHLIAAMWLQEGILARRNGNGOIEPLGKADVAAGQFVLEILRRANRPA AEDAPPTLQWRVYRNMAQSGAPLITITPEQKORQSGKMTGRVREPLIDNLLT DPLSKIASLTSVPHGVHDNYALFELRLMILGHNAMDDONALPVEVFCQSLIT LQVVEVQSKKPKASKQTDGASKKKAAATRRSSKRSKGSDELIGNATATMT PESELDSEAPSDGLIPDDLIIPDEKVEVQFOPRLSYDGLIDIVQKEIIVGASPV EPALILADADBRIVLCTLADPAALIGILOFLRGDYRLDILHAATASRLAAGSIDS LIRVELPQLAGPIVAPCELVQWRPARGAKLKLVSYSRLSDSLIPGAPLIV CLARGVRLQDRDKALSESAAKTYORGSRLSHESTMEVVALTDHAAIDLGLHD AGDESPIVPEGETTFVRVREGITPKSLIKIKDDSDPMQGTGTVTVGDVPLQDLM AVREDRQLVQVDRPAKISAFRRRLVQGLDITLVSEKQQLRAGAAVPLADVLGD VTLLETTAAMHQAIERLDSLOWSPPRPEGLDITLVQDGRMTARLSRMVGGVLA DDMGKGTVOGLIIVERGVPVGPALVAVPVGSEWVREAEKFTPLTGKLRDGD EVITKAGPNDLIVSYOLLORDAKRFAPRGWHTLVLDRAOPIKNSQTSOAIPTIOA DMRLGSGTPIENHGLMSLFRILSPGLLSMQORFRFRAPPIRHSDDARRSLAR LVRFVLRKTKVKEPLPRTETILRAELSAPEKLYEDARLALAEITMGGAASH EGRRRIQTLISWLRRLQLACHPSLVPEPMKGTSSKLQPLSLVEELRGGDRALVSO FVKHLSVVRALDERGISYQYLDGATPSHEKQORRDARQNGEGDIFLILKAGTGILN LTAADYVHLDPMMNPVAVEDQATDBRAHIGERAVTVYRLVARTIEQIILQHADKR ELVAGILGTBAAAKLOTODLIDLKAMG"			
complement(5458..5850)			
/locus_tag="RB13"			
complement(5458..5850)			
/locus_tag="RB13"			
/codon_start=1			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="CAD71335.1"			
/db_xref="GI:32442817"			
/translation="MHLPHGNQRFQRRTKMYRETSETLLRTLATETVRLTPVAPHP PRRSFELMETPESSDPQSLKDCPKIPSGTIPGINKVSCSVIPEKLPHTTCQEPPLSNR AGNRVVDLKRGVWFGTRPVRLRLSP"			
5836..7386			
/locus_tag="RB14"			
5836..7386			
/locus_tag="RB14"			
/codon_start=1			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="CAD71336.1"			
/db_xref="GI:32442818"			
/translation="MGQMFDPVPAANEFFESTREWADYLCGIEGVPQMSRNSNGR FTLTGCVDTSAKLEFLTCPIPSIGYROLSTSLRCIDRSRLFVEILARSCYVRVQAD VMORGALVLSRFDPLTRGTRFLDAQCPADPSCVDALEALAEILAEHSNLEGLD FSTOSVARRSRRETRLSMLAVGLPRLPADEVIRADMGSGASNSIDQATTTAFNA AAVRLSMGEIETDSGRVRYEPTDEALTAQMGRLAVYGVYIDRYKGLPMLTLLDD DFDKLMSWTWTSVEKTVQFRGRVNLMAKGVNAGPLGFDDEGVWFFSIGILTIR KCDPPTVLMGSLIDQPCGEYIARDEQGSPLHFAVALRSGLAGIQLNFRGPDGD THPSAVFGCMIDRMGTILNAPLWQVLSVAGAPGDAARAPITQRPISPEINGSDPA WAKOIEFVQPLIEITLISKQIVHVAVMDGMSDQHPHMTPHAGVIDAKGRPRMLAYLT KIREMLI"			

```

Db      184884 GAAGACTGACGTCACTTTGAAACCAAGTCGAAAGATCCCTCTCGACAAAGTC 184943
Qy      259 GTGAAATGGCGAAATGGACCGGCTGACATGCTGGCCGCGCGCTTAAAGCGCCAGCGTT 318
Db      184944 ATCAATGGCGAACTTCACCGCGGACACATGTTGGCCGCGCAACCGATCGTCACAGATC 185003
Qy      319 GGCTTGGGACCAAGGCGGATATCGAGGTGCGGTTCAGCAATGGGCTTCAACGGCAACGGG 378
Db      185004 GGTGTGGAAACCAAGTCGCCAATCTGACCGCTTACGCTTGACCTGACATCAACGGCAATGGC 185063
Qy      379 ATTACGGTGTCCATGAAATCTGGCATCAGATGAGCCGCAACATACC---GCTGAAGGC 435
Db      185064 ATCAGCGGTAGCAACGAAGTTTGGGCAAGATGACAGAGAACATCTGAACTGAAGAGC 185123
Qy      436 GGTAAACGGGTGATCCGATCAAGGCAATTTCTTAAACCGGTGCTCGAAAAATAC--- 492
Db      185124 CCCACACCGAAACATCCCATACAGCGCGGACGCTGAAACCGATTCGAGGATTACCTG 185183
Qy      493 AAAGCCGAAGGCAAGCGGCTTCAATATGGGATGACCTTCCCGCGGATCTCAACATC 552
Db      185184 CAAAGACGGGGCGAACCATTTCCGATGGGATGATGTTCTGTTAGCACTCAAAATTAC 185243
Qy      553 AAATGCGTTATGCTGCTGGCGGCTGGGATCAATCTGCTACTATTGCGCCGCGCAG 612
Db      185244 GAGATTCGATATTGGCTGGCGCGGTTCATCCGCGCATGACACCGAATC---T 185300
Qy      613 GACATTTCCGGCCAAATCGCGCGAAGCCTTGTGTGCTGATGCCCGCGCGCAATG 672
Db      185301 GACATCAAAAGGCTTCAACGACGACAGTCAATTTGTCAACGATGTCACGCTCAAAATG 185360
Qy      673 CCGTCACGCTGGAAACCGGACCATTTTGGTATTTGGTGGGCGAGCGGAGAACGAG 732
Db      185361 CCTCAAGACCTGGAAACGGGACATCTCAAAAGTTACTGGGTTGCGCAACCGTGAACCA 185420
Qy      733 CAGCGCGTTCAGAGGCAATAGCGCTGCGGTGATCAGCATGAAGAACTTGAAGAC 792
Db      185421 AAAGCGGTGTGATCGATCGGTGTTCCAGTACACAGCACTACACATCTGGAAGAC 185480
Qy      793 ACCCGGAAAAAGTCTTGGCGGTGACCAAAACAATGGCGGAGAAATACCCCAACACTAT 852
Db      185481 AACCCGAAAAAGTCTTGGCGGTGACCGGAGATGGCGGAGAAATACCAAACTCAC 185540
Qy      853 CTGGCGGTGACCAAGGCGGTGATTTGGGCGCGGATGCTGGAGCGCGCAATTAACAG 912
Db      185541 TTGGCGGTGATCAAGGCTGTGATCCGTGCGCAAAATGCTGATGCCACGATGACTCG 185600
Qy      913 -----AACCGCAAGAGCCATCGAAATCTGGCGGCAAAAACAATACGTCGT 960
Db      185601 GGCAAATTGTCATATCGCAAGAGCCGTGGAATCTTCAAGCCGCAAAATTAATGTCGT 185660
Qy      961 GCTGACGTGAAGTGTCTGGCGGCTAGCATGAACGGCACTTTGCAATACGAAAAAGCAT 1020
Db      185661 GCGGAGAAAGAGGTGATGGAACAATCGATGATGGGACGTTTGTCTTCAATTCACGAC 185720
Qy      1021 AAACGGCGGCTACCGGACCTTCAACACTTTCGCGACGCGCGCAAGCTATCCGTCTAC 1080
Db      185721 GTTTCGGAATGCGGACCTTCAACGTTCTTCAACAGAAACGACGATCCGCACTAC 185780
Qy      1081 ACGAGTGAAGTGTGATCTGACCCAGCTGAGCGCTGGGCGCATGATCAATGAATTCAA 1140
Db      185781 ACGGATGGAATTTGGTTCTCTCAACAGATGCGACGCTGGGGAACAATCAAGATTCAA 185840
Qy      1141 CCGGACAACTGATATCTGATACCCGCAAGAAAGTCTTACCGCGCGGACATCTATCTGCGC 1200
Db      185841 CCCGCGAGTTGGATCCCGCAAAACGCGCAAGAAAGATCTACAGCTGATCTATCGCCA 185900
Qy      1201 GCGGCGAAGAACTGGTGCAGAGGCAAGCGCAAGCGGAGACTTCCCTCTC 1253
Db      185901 GCGCGGAGCTTCTAATACGCAAGGCAAACTGATCCAAATGAATCCGCGC 185953

```

RESULT 6
AE017282_06/c

WPCOMMENT

Sequence split into 33 fragments. LOCUS AE017282 Accession AE017282

Fragment Name	Begin	End
AE017282_00	1	110000
AE017282_01	100001	210000
AE017282_02	200001	310000
AE017282_03	300001	410000
AE017282_04	400001	510000
AE017282_05	500001	610000
AE017282_06	600001	710000
AE017282_07	700001	810000
AE017282_08	800001	910000
AE017282_09	900001	1010000
AE017282_10	1000001	1110000
AE017282_11	1100001	1210000
AE017282_12	1200001	1310000
AE017282_13	1300001	1410000
AE017282_14	1400001	1510000
AE017282_15	1500001	1610000
AE017282_16	1600001	1710000
AE017282_17	1700001	1810000
AE017282_18	1800001	1910000
AE017282_19	1900001	2010000
AE017282_20	2000001	2110000
AE017282_21	2100001	2210000
AE017282_22	2200001	2310000
AE017282_23	2300001	2410000
AE017282_24	2400001	2510000
AE017282_25	2500001	2610000
AE017282_26	2600001	2710000
AE017282_27	2700001	2810000
AE017282_28	2800001	2910000
AE017282_29	2900001	3010000
AE017282_30	3000001	3110000
AE017282_31	3100001	3210000
AE017282_32	3200001	3304554

Continuation (7 of 33) of AE017282 from base 600001 (AE017282 Methylococcus capsulatus

Query Match 12.3%; Score 171.2; DB 1; Length 110000;

Best Local Similarity 53.3%; Pred. No. 8.4e-23;

Matches 542; Conservative 0; Mismatches 408; Indels 66; Gaps 6;

```

Qy      125 TGAATTCGCTTCAATCAAGCTCACCGACATGACCGCCGCTGGCGGCGGCAAAAG 184
Db      18318 TGAAGTGGGTTTCAATCTCTTACAGATGCGGTCTCCGTGATCTGCGGAAAAAG 18259
Qy      185 GCTTCTGAGGACAGGCGCTGTTGTCGATCTGAAAGCGAGGCAACTGGAAGGTGG 244
Db      18258 GTTACTTCCGCGACACGCGCTGACAGTCACTGTCCGCGCAAGCCGCTCTGGCGGCA 18199
Qy      245 TGATGATAGGGTCTGTGAATGCGAATGACGAGCTCGCACTGCTGGCGCGCGCGCT 304
Db      18198 TCCGGAAGAGTTCGACACCGGCTGTCGAGCGGCGCCACATGCTCGGCGGATGCCGA 18139
Qy      305 TAGCGGCGGAGCTTGGCTT---CGGACCAAGCGCGGATTCGAGGTGCCGCGCT 361
Db      18138 TCGCCACCACTCTCGGCTGGGGGGGGTACGGAAGGCTACATTCGCGCGATGTGCTGG 18079
Qy      362 GCTTCAACGGCAACGCAATTAAGTGTCCAAATGAATCTGGCATGAGATGAAGCGGAACA 421
Db      18078 ATCTGAATGGAACCGCGTGAAGGTGTCCAAAGCCCTGTACC-----GGCGCTCA 18028
Qy      422 TACCGGTGAAAGCGGTAAACCGGTGATCCGATCAAGGCAAGCATTAATCAAAACGGTGG 481
Db      18027 CGGCGGTGGCGGCTGACGCTGCGACGCGCGCGGTTTCGCGCGCTGCTCGCGGCTT 17968
Qy      482 TCGAAAAATACAAACCGCAAGCAAGCCGTTCAATATGCGCATGACCTTCCGCGCGGAT 541
Db      17967 TGTCTGACAGGAGCGGAGGCTCTGCGGTTTCGCAAGGTATTTCCCTTTCTCC-----T 17914
Qy      542 CTCAACAATCAAACTGCGTTTACTGCTGCGGCGGTGCGGATCAATCTGGCTACTATT 601
Db      17913 CGCACAATCTACTGCTGCTGCTATTGGCTGGCGTGGCGCGGCGGTGATCC----- 17865

```

QY 602 CGCCGCCGAGGACATTTCGGCAATGCGCGACAGACGCTTGTGTGCTGACCCGCG 661
 Db 17864 -----CGACCGCCACGTCGCGCATGACGTCATTTCGC 17833
 QY 662 CGCCGCAATGCGCTTCACGCTGGAAGCGCGACCATTTTCGTTATTCGTCGCGGAGC 721
 Db 17832 CGCCGAGATGCAAGTTCCTGCGCGCGGGAGATTTCCGCTACTCCGCGGGTGAAC 17773
 QY 722 CGTGAACACGACGCGCTGTTCAAGGCGATAGGCTGCCGGTGATCACCGATGAAGAAC 781
 Db 17772 CCGTGAATCCCGTGCCTGCGCCGAGGAGATCGCGCTACTGATCACCGGTCAATGA 17713
 QY 782 TCTGGAAGGACGCGCGGAAAAAGTTTGGGGGTGACCAATCATGGGGGAGAAATACC 841
 Db 17712 TCTGGAACACAGTCCCGAAAAAGTCTGAGCGCTACACAGGACTGAGCGCAACCATC 17653
 QY 842 CCAACACCTATCTGCGGCTGACCAAGCGCTGATTCGGGCGCGCATCTGCGACGCGC 901
 Db 17652 CCAATACCTTGGGAGCGCTCTCATGCGCTGCTGGAACCGCGCAATGGCT-----CG 17539
 QY 902 ACAATTAACAGAACCGCAAGAACCCATGAAATGCTGCGCAAAAAACATACGTCGCTG 961
 Db 17539 ATCGCGCCGAAAAACCGCGGAAAGCGCGCCCTCCTCGCAATCCGCACTATGTCGCGC 17539
 QY 962 CTGACGTGGAAGTCTGCGGCGCTAGCATGAAGCGCACTTTCGATTCGAAAAAGACGATA 1021
 Db 17538 TCGCGCGGAGATCTTCGAGATGATGATGATGATGATGATGATGATGATGATGATG 17479
 QY 1022 AACGCG---CGCTACCGGACTTCAACACCTTTCGCGCACGCGCGCATATCCGCTCT 1078
 Db 17478 CGGCGGTGTCCTCCCGGACTTCAACGATGATGATGATGATGATGATGATGATGATG 17419
 QY 1079 AAGACAGTCACTGCTGATTCGACCTGAGCGCTGCGGCGCATGATCAATGAA 1134
 Db 17418 GTTCCACGCGCTGTGTCTTGAACCCAGATGCTGCGGCGGAGATCGCCGAA 17363

RESULT 7
 AE008293 10676 bp DNA linear BCT 28-MAY-2004
 LOCUS Agrobacterium tumefaciens str. C58
 AGRI008293 AE007870
 ACCESSION 187 of the complete sequence.
 VERSION AE008293.1 GI:15159385
 KEYWORDS
 SOURCE Agrobacterium tumefaciens str. C58
 ORGANISM Agrobacterium tumefaciens str. C58
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 REFERENCE 1 (bases 1 to 10676)
 AUTHORS Hinkle,G., Slater,S.C. and Goodner,B.
 TITLE Complete Genome Sequence of Agrobacterium tumefaciens C58
 (Rhizobium radiobacter C58), the Causative Agent of Crown Gall
 Disease in Plants
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 10676)
 TITLE Hinkle,G., Slater,S.C. and Goodner,B.
 JOURNAL Direct Submission
 SUBMITTED (14-AUG-2001) Bioinformatics, Cereon Genomics, 45 Sidney
 Street, Cambridge, MA 02139, USA
 COMMENT Approximately 800 bp of telomeric sequence missing from the left
 end of the chromosome and 200 bp missing from the right end.
 FEATURES
 Location/Qualifiers
 1..10676
 /organism="Agrobacterium tumefaciens str. C58"
 /mol_type="genomic DNA"
 /strain="C58"
 /isolate="Cereon"
 /db_xref="taxon:176299"
 complement(93..1352)
 /gene="AGR_L_1874"
 complement(93..1352)

/gene="AGR_L_1874"
 /note="HYPOTHEICAL 40.1 KDA PROTEIN IN SIPU-KIP1
 INTERGENIC REGION"
 /codon_start=1
 /transl_table=11
 /product="AGR_L_1874p"
 /protein_id="AAK89510.1"
 /db_xref="GI:15159386"
 /translation="MERKSMEOKKLTPTADLRKSRASLGAIFMAATSAIPGPIF
 QTAIFITDUGAPAFGLIATLIDFVQDNTMRVTLTRKASDILANAFIPAGIFLA
 ILVITIGLFENIGNGTGLNLAFGLDPIKGAISAFISIAIFVSRAGLAVDFI
 IFAGILMTVLTVYAFVSPAPVGDARFQTPDPIINFATITITVIGTGGYITVGAH
 RLDRGMVGINLPAVNRALTLGIAVTGIMRVFLAILGVASGVITDTSQANPA
 AQRSAAGDIGPRLFGIIFMAAITSVGAATSVSPFPVKODMSRRAMVAVF
 IAVSVCTLTTPPAMLVFVGLNGILFISIFLPAWKEDLMGHRYSILL
 GLGVLTCLTIYMWGKSRGTLFGLGL"
 1517..2266
 /gene="AGR_L_1876"
 1517..2266
 /gene="AGR_L_1876p"
 /note="(AUJ3213) putative GntR-family transcriptional
 regulator."
 /codon_start=1
 /transl_table=11
 /product="AGR_L_1876p"
 /protein_id="AAK89511.1"
 /db_xref="GI:15159387"
 /translation="MCYQGMPCRMPTPEPMLHCSSERAMTETIGLPLADRLAK
 GIRTLISGELAPQORLEAFAFSRFDPSRSLLEAFLLTKDGLRREARNGVAV
 PTMSVIDIIVYRMBVEQALQAMHKIPAKMRAVEDAKIRREADMDVGVSAM
 KEHAAYVDLQDSARISDFYERIAELRLCFGLKDPEQLHSPFVDMNSKISLMEEG
 AOBAGNEMELVNLISERTVLCALRAVHTS"
 2341..2955
 /gene="AGR_L_1879"
 2341..2955
 /gene="AGR_L_1879p"
 /note="response regulator nasr - Azobacter vinelandii"
 /codon_start=1
 /transl_table=11
 /product="AGR_L_1879p"
 /protein_id="AAK89512.1"
 /db_xref="GI:15159388"
 /translation="MLRCVINADMTLRDLITLVIDENARASIIIEGREGAYQRTV
 IHEVNGVARTIETLOPYIFIDLPNPNDDMHELFOLRTVGRPIAMVVDSDTASIS
 AAVAGYAAVYVDLTKKRVKPIIDMAVSRNARSRLQRELAQSALEBKIVIERAK
 GILMKRGLSEEBEPALIRQTPANMEKKKISIAQSVTAAGILM"
 2991..4286
 /gene="AGR_L_1881p"
 2991..4286
 /gene="AGR_L_1881p"
 /note="hypothetical protein"
 /codon_start=1
 /transl_table=11
 /product="AGR_L_1881p"
 /protein_id="AAK89513.1"
 /db_xref="GI:15159389"
 /translation="MAAPEKTAGSKGEATQWRDAPALVGSROKILRAGPIBIVASV
 LIAAERFADREGLIDLVQDVSWANVRDLAROPPIAMLSMPVPSNLGGSNP
 SPITTPSLRGSNALITLSTRLEAFMRKALTLSTAGLLENARALKVLDMRAGGA
 PPTIGMTYPPFSNHYEFKRYWLAAGIHFDHDKLVVPPPTSDALAGAIDFGVGA
 PWNIVAAERGVIRVAAKODLWPSAPKEVIGMRBEMASQOETVGRILITLADAAWGC
 DLANDHDLSCALADPRYIGAPQSIIRVLVAGERSIDSGNRVIEKFTFHGDHANY
 PROSOLIMIVSOALRMGQAEISFGVNAALSAVRPDYRALGSGKAGADDIRIEGQ
 DEGRFVDFGVFPDADTAGVNSPAVRTSVVSAPSGEA"
 4544..5959
 /gene="AGR_L_1886"
 4544..5959
 /gene="AGR_L_1886p"
 /note="nitrate-binding protein ntrA precursor, periplasmic
 - osillatocicean cyanobacterium"
 /codon_start=1
 /transl_table=11
 /product="AGR_L_1886p"

gene
/protein_id="AAK89514.1"
/db_xref="GI:15159390"
/translation="MLPGTTPVKAQTSGVFFLIRPRKRHRHPKMLNKKNNRMKK
ITSGVSRRTILIKTTATALVTAATAVPSSGAFAATAPBEVAKLGFALTDAAPLI
IAEKLFPKMGPEVEVLKQASMGATNGLVGLASGIDGAILTPMPYIMTKGV
TNNPVVPTIILARLNLDQGISVAKEVAEAGVOLDASKLAAEPCKAOKSEIKAM
PGGVAHWIRMYLAAGGIDDDPKDVSITVPPQPMVNMVKGMDVPPKSGPBNBD
VNOGIGFTACTTGELMKHPKALGMRADWYKPNATKALLAAMEAQOCDBMARK
EENSTILIGKQFNVPVKDVLGRKLKNGNGRVALENTGQMPFQWDHASPYSHD
SMPTIENKMGFPADTPVKALVAKNREDIWRRAAKDLGVADLPASTSRKETFPDGG
KYPDENPSPYLESISIKAS"
6067..6966
/gene="AGR_L_1888"
6067..6966
/gene="AGR_L_1888"
/note="NITRATE TRANSPORT PERMEASE PROTEIN NRTB"
/codon_start=1
/transl_table=11
/product="AGR_L_1888p"
/protein_id="AAK89515.1"
/db_xref="GI:15159391"
/translation="MPMSLARKFNBBPQOPAVOKANVALSPHLPRIDLRKKWGA
AANTPIPLAVLAVGQMLCSAGSSLPSPQVFEESYDLIVSPFNHSGODIGL
LAVLSLEVAVAGPGLAALAGILGAVIGSVVMARGDLPYOLRTKVPPLAPLAKI
LAPQDSNPBAITFVITITSITWPIITANGVRNIPDRIRNAKVLRLNQDEFFKINLP
SAAPYITGLRIGLISMLAIVAAEMLTGVGIGLFIWDANSSRLPDIIIVALYIGV
VPMFDKLVAAAGKITGGAAN"
6980..7777
/gene="AGR_L_1890"
6980..7777
/gene="AGR_L_1890"
/note="NITRATE TRANSPORT PROTEIN NASD"
/codon_start=1
/transl_table=11
/product="AGR_L_1890p"
/protein_id="AAK89516.1"
/db_xref="GI:15159392"
/translation="NNAYILKLDHIDKPYDKGSGRAVVLKGINLTIEKEEFSVYIGSG
CKSTMLNLIAGLITVSAVAVLENREVENPESRAVVPQHSLLPMLSYENNVLAIV
AIVPNTKTRFRHDWNRNLDLVQMAHXDKQKSEISGGKMGVLAARLAMEPKIL
ILDEPGALDLATRAHLODAVMEIHARGNMTIHDVPAVLISRLITMTNGPAA
HIGEVLDPLARPRRIELASDKITLYKREAVLFLYERHFEVAE"
7714..10230
/gene="AGR_L_1891"
7714..10230
/gene="AGR_L_1891"
/note="asemilatory nitrite reductase large subunit PA1781
(imported) - Pseudomonas aeruginosa (strain PAO1)"
/codon_start=1
/transl_table=11
/product="AGR_L_1891p"
/protein_id="AAK89517.1"
/db_xref="GI:15159393"
/translation="WPGCAEVPRLTSPLRSGVITMTQKLVITNGMAPRMLNLT
FTAGLGDVITFNAEPVNYDRIMLSVLSEKSYEDIVIHNDWYMANNVLLHKA
KTRIDRANKITSENGITASVYDLVATGSLPTIIPVPGHQLRGVLAIVLDDVTMG
LEIAGKGAQVIYAGLIGLEAVGLKROGMDTVIHLMPITEROOLDPAAVYLEKA
INERGIIDITTKANTCIIIGKEKVBGIELDERRIIKGMVWVANGVIRPAGSLAENAGA
VNRGIVDDGMNTSPASITYALGECBAHRKCKYGIYVAPLYESARLATDLRCGSNAETHG
SVTNKFLKKTGINTFSADCFABDDREIIVLRDVTAGVYKRLIKENRIIGAVLYGET
ADGSPFLDKMGSTDISAMRETLFGQAVQGSPLDPMAAVALPDAEIGCGNGVC
GKITATYSLKGLTSLDVRATKASACGCTGVEQLMTLGDITVNPAAVOPMCKC
TDLGHDVRLIKAGKGLTPIVAMQELBETSCGCAORPALANYILVCDPDEFRADY
OSRYINERHNVAKIDGTVGSVPPRMGGVTNADRLRALIVDVKPEIPLKYTGCGRI
DLTGLEKEDLVANWADLKAGPISGQATKQALRYTKYTCVGEQKRFETGOSTGJREI
EKFPMGSPYPAKLVASGCPNCKEATVCKQIGVIVDSGELHFPAAAGDILKGEV
IGLVHTEDEAHEHIALVATQMYREQKRYLERIYKAKSIGDYEVARQITMDADRKAYF
DRFVSEKFAQVDPSERSVSGHDKKEFPMAAIGFSEAAE"
10227..10571
/gene="AGR_L_1892"
10227..10571
/gene="AGR_L_1892"
/note="asemilatory nitrite reductase small subunit PA1780
/protein_id="AAK89518.1"
/db_xref="GI:15159394"
/translation="MLPGTTPVKAQTSGVFFLIRPRKRHRHPKMLNKKNNRMKK
ITSGVSRRTILIKTTATALVTAATAVPSSGAFAATAPBEVAKLGFALTDAAPLI
IAEKLFPKMGPEVEVLKQASMGATNGLVGLASGIDGAILTPMPYIMTKGV
TNNPVVPTIILARLNLDQGISVAKEVAEAGVOLDASKLAAEPCKAOKSEIKAM
PGGVAHWIRMYLAAGGIDDDPKDVSITVPPQPMVNMVKGMDVPPKSGPBNBD
VNOGIGFTACTTGELMKHPKALGMRADWYKPNATKALLAAMEAQOCDBMARK
EENSTILIGKQFNVPVKDVLGRKLKNGNGRVALENTGQMPFQWDHASPYSHD
SMPTIENKMGFPADTPVKALVAKNREDIWRRAAKDLGVADLPASTSRKETFPDGG
KYPDENPSPYLESISIKAS"
6067..6966
/gene="AGR_L_1888"
6067..6966
/gene="AGR_L_1888"
/note="NITRATE TRANSPORT PERMEASE PROTEIN NRTB"
/codon_start=1
/transl_table=11
/product="AGR_L_1888p"
/protein_id="AAK89515.1"
/db_xref="GI:15159391"
/translation="MPMSLARKFNBBPQOPAVOKANVALSPHLPRIDLRKKWGA
AANTPIPLAVLAVGQMLCSAGSSLPSPQVFEESYDLIVSPFNHSGODIGL
LAVLSLEVAVAGPGLAALAGILGAVIGSVVMARGDLPYOLRTKVPPLAPLAKI
LAPQDSNPBAITFVITITSITWPIITANGVRNIPDRIRNAKVLRLNQDEFFKINLP
SAAPYITGLRIGLISMLAIVAAEMLTGVGIGLFIWDANSSRLPDIIIVALYIGV
VPMFDKLVAAAGKITGGAAN"
6980..7777
/gene="AGR_L_1890"
6980..7777
/gene="AGR_L_1890"
/note="NITRATE TRANSPORT PROTEIN NASD"
/codon_start=1
/transl_table=11
/product="AGR_L_1890p"
/protein_id="AAK89516.1"
/db_xref="GI:15159392"
/translation="NNAYILKLDHIDKPYDKGSGRAVVLKGINLTIEKEEFSVYIGSG
CKSTMLNLIAGLITVSAVAVLENREVENPESRAVVPQHSLLPMLSYENNVLAIV
AIVPNTKTRFRHDWNRNLDLVQMAHXDKQKSEISGGKMGVLAARLAMEPKIL
ILDEPGALDLATRAHLODAVMEIHARGNMTIHDVPAVLISRLITMTNGPAA
HIGEVLDPLARPRRIELASDKITLYKREAVLFLYERHFEVAE"
7714..10230
/gene="AGR_L_1891"
7714..10230
/gene="AGR_L_1891"
/note="asemilatory nitrite reductase large subunit PA1781
(imported) - Pseudomonas aeruginosa (strain PAO1)"
/codon_start=1
/transl_table=11
/product="AGR_L_1891p"
/protein_id="AAK89517.1"
/db_xref="GI:15159393"
/translation="WPGCAEVPRLTSPLRSGVITMTQKLVITNGMAPRMLNLT
FTAGLGDVITFNAEPVNYDRIMLSVLSEKSYEDIVIHNDWYMANNVLLHKA
KTRIDRANKITSENGITASVYDLVATGSLPTIIPVPGHQLRGVLAIVLDDVTMG
LEIAGKGAQVIYAGLIGLEAVGLKROGMDTVIHLMPITEROOLDPAAVYLEKA
INERGIIDITTKANTCIIIGKEKVBGIELDERRIIKGMVWVANGVIRPAGSLAENAGA
VNRGIVDDGMNTSPASITYALGECBAHRKCKYGIYVAPLYESARLATDLRCGSNAETHG
SVTNKFLKKTGINTFSADCFABDDREIIVLRDVTAGVYKRLIKENRIIGAVLYGET
ADGSPFLDKMGSTDISAMRETLFGQAVQGSPLDPMAAVALPDAEIGCGNGVC
GKITATYSLKGLTSLDVRATKASACGCTGVEQLMTLGDITVNPAAVOPMCKC
TDLGHDVRLIKAGKGLTPIVAMQELBETSCGCAORPALANYILVCDPDEFRADY
OSRYINERHNVAKIDGTVGSVPPRMGGVTNADRLRALIVDVKPEIPLKYTGCGRI
DLTGLEKEDLVANWADLKAGPISGQATKQALRYTKYTCVGEQKRFETGOSTGJREI
EKFPMGSPYPAKLVASGCPNCKEATVCKQIGVIVDSGELHFPAAAGDILKGEV
IGLVHTEDEAHEHIALVATQMYREQKRYLERIYKAKSIGDYEVARQITMDADRKAYF
DRFVSEKFAQVDPSERSVSGHDKKEFPMAAIGFSEAAE"
10227..10571
/gene="AGR_L_1892"
10227..10571
/gene="AGR_L_1892"
/note="asemilatory nitrite reductase small subunit PA1780

```
{imported} - Pseudomonas aeruginosa (strain PAO1)"
/codon_start=1
/transfer_table=11
/product="AGR_L_1892p"
/protein_id="PAK9518.1"
/db_xref="GI:15159394"
/translation="WRMSDNIWDIGISDIPLRGARCVTPNGKIAVFTAEANOVA
IEDHCPRKGGPLSGIVHGASVTCPLNNWVSIETGALGADGEVETIPVRINDGRL
SIALETTIMAAE"
```

Query	March	9.6%	Score 133;	DB 1;	Length 10676;
Best Local Similarity	50.2%	Pred. No. 1.9e-15;			
Matches 572;	Conservative 0;	Mismatches 495;	Indels 72;	Gaps 7;	
QY	CGATGTGGGGCACTTGGCAAGCTGCGAAGAAAGAAAGATTGTAATTCGGCTTCATCAAGCT	146			
Db	CGCTTTGGCGGCCACCGCAGAACCGGAAGTGAAGGGTGCAGAAACTGGTTTCATCGGGCT	4827			
QY	147 CACCGACATGGCGCGCCGTGGCGGTGGCCCGCGAAGAAAGCTTCTTGACAGACGAGGCTT	206			
Db	4828 CACCGAGCGCGCGCGCTCATTCATTTGGCGCGCGAAGAGGACTTTTCCCAAGCATGTCAT	4887			
QY	207 GTTGTCGAACATGGAAGG---CAGGCGCACTGGAAGTGTGTATGTATGGGTCCGGA	263			
Db	4888 GCTTAGGTGAGGTTCGTAAGACAGGGCTCTCGGGGGGCGACGCGCGACATTCGTGCT	4947			
QY	264 TGGCGAACATGACCGGCTCGACATGCTGGCGCGCGCGCTTATGCGGCGAGCTTGCTT	323			
Db	4948 CGGGCG-----CGCTTCGAACCGCATTCGATGGCGGCGCATTCCTGACGCGCATGCTTA	5001			
QY	324 CGGCAACCAAGGCGCATTCGAGGTGCCGTTTCAGCATGGGCTTCAACGCGCAACCGATTAC	383			
Db	5002 TCTTATGCATACCGGCAAGGTGACGCGAACAACGTCCGGTCCGATGACATCTCTGCG	5061			
QY	384 GGTGTCCAAATCTGGCATCAGATGAAAGCGCAATACCGCTGGAAGCGGTAAAC	443			
Db	5062 ACGCTTAATCTGCACAGCGCAAGGCGAT-----TTCCGTGCGCAAGGAATATGACAGAC	5115			
QY	444 GTTGCAATCCGATCAAGCGAGATTATCTTAAACCGGTCTCGAATAATCAAGCCGAAAG	503			
Db	5116 CGGGTCAGATCGATGCTCTTCAAGCTGAAGGGCGGCTTTGAGAAAAAGAGCGACGG	5175			
QY	504 CAAGCGCTCAATATGGCGATGACCTTCCCGCGCGATTCACAAATCAAATCTGCTTA	563			
Db	5176 CAAGGAATCAAGGCGCGCATGACCTTCCCGGCGGACGCGATCTTCGATCTCGCTA	5235			
QY	564 CTGGCTGGCGCTGGCGGTATCAATCTGGCTACTATTCGCGCGGACAGACATTTCCG	623			
Db	5236 TTGGCTGGCTGGTGTGGCATTCATCCGACAG--	5269			
QY	624 CCAATCGGCGCAGACGCTTGTGTGGTGATGCCCGCGCGCAATGCGGTCCACGCT	683			
Db	5270 -----GACGCTTCCACCATCGTCTGTTCCGCGCCCAATGCTCCCAACAT	5316			
QY	684 GGAAGCGCGACCAATTTTCGTTATTCGCTTCGCGGAGCGCTTGAACCAAGAGCGCGGTT	743			
Db	5317 GAAAGTGGCAACATGACGCTTCTGTGTGAAGGAGCGCTGGAACAGACGCTCTGTTAA	5376			
QY	744 CAAGGGCATAGGCGTGGCGGTGATTCACCGATGGAAGACTTGGAAAGGACACGCGGAAAA	803			
Db	5377 TCAGGGATCGGCTTTATCTGCTGTACACCGGGGAATCTTGAAGAGGATCCGAAAA	5436			
QY	804 AGTCTTGGCGTGAACCAACATATGGCGGAGAAATACCCCAACCTATTTGGCGGTGAC	863			
Db	5437 GGGCGCTGGGATGGCGCGCCACCTCGGTGGGAAAAAGAACCCCAAGCGCACCAAGGCTTTGCT	5496			
QY	864 CAAGGGCTGATTCGCGCGCGCGATTCGGCTGGACGCGCGCAATTAACAAGAACCGCAAGGA	923			
Db	5497 GATGGCGGTATGAGGCGCAGCAATGGTCCGACAGATGGCGCAACAAGAGAGATGTC	5556			
QY	924 AGCATCGAAATGTGGCGCAAAAACAATACGTGGTCTGACGTGGAATGCTGGCGGC	983			

Accession	Sequence	Position
Db	5557 CACCATC-----CTCGGCAAAACGCCCACTGGTTCAACGTGCCGCCGAAGATGTTCTCGG	5610
Qy	984 TAGCATGAACGGCACTTTTCGAAATACGAAAAAGACGATAAACGCGCGCTACCGGACTT---	1040
Db	5611 CCGCTTCAAGGGCAATATTAATTACGGCAACGGCCGGGTGCTGAAATAATCTGGGCTCA	5670
Qy	1041 CAACACCTTTCTTTCGCCACCGCGCCAGGTATCCGCTCTACAGCACTGACGTCTGTATCT	1100
Db	5671 GATGAAGTTCTCGCAGCAACATGCTCTATCTGTTCCACAGCCATGACAGCTGGTTGAT	5730
Qy	1101 GACCCAGCTGAGCGCGCTGGGCATGATCAATGAATTTCAACCGGCAACTGTATCTTGA	1160
Db	5731 CACGGAATAATCCGCTGGGGAAAAATTCGCGCGGTGATACGACGTAAAGCG	5781
Qy	1161 TACCGCCAAAGACGTCTACCGCCCGGACATCTATTCGCGCGGCGCAAGAACTGTGTC	1219
Db	5782 GCTGTTGTCGAAGGTGAACCGGAGAGATATCTGGGCGCACGCGGGAAGATCTTCGGC	5840

RESULT 8	AE009321/c	10707 bp	DNA linear	BCT 28-MAY-2004
LOCUS	AE009321			
DEFINITION	Agrobacterium tumefaciens str. C58 linear chromosome, section 91 of 187 of the complete sequence.			

ACCESSION	AE009321	AE008689
VERSION	AE009321.1	GI:17742339

KEYWORDS	SOURCE ORGANISM
Agrobacterium tumefaciens str. C58	Agrobacterium tumefaciens str. C58

REFERENCE
AUTHORS

Wood, D.W., Setuhal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.
1 (bases 1 to 10707)

Rhizobium/Agroacterium group; Agrobacterium.

Rhizobiaceae; Rhizobium; Agrobacterium group; Agrobacterium.

Bacteria; Proteobacteria; Alphaproteobacteria; Rhiclobiales;

Chen Y., Moo L., Kitchajima J.P., Okura V.K., Almeida J., N.F., Zhou Y., Boyce Sr., D., Chapman, P., Clendenning J., Deatherage G., Gillet, W., Grant, C., Genttner, D., Kutyavlin, I., Levy, R., Li, M., McCelland, E., Palmeri, A., Raymond, C., Rouse, G., Saenphitumachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Iiao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tinney, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Neeter, E.W.

The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58

JOURNAL
Science 294 (5550), 2317-2323 (2001)

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 21608550
 11743193
 2 (bases 1 to 10707)
 Chen, Y., Setuwal, J. C., Kaul, R., Monks, D., Chen, L., Wood, G. E.,
 Wood, Y., Woo, L., Kitajima, J. P., Okura, V. K., Almeida Jr, N. F.,
 Zhou, Y., Boyce Sr, D., Chapman, P., Cledenning, J., Deatherage, G.,
 Gillet, W., Grant, C., Guenther, D., Kuttyavin, T., Levy, R., Li, M.,
 McClelland, E., Palmieri, A., Raymond, C., Rouse, G.,
 Saenphimachak, C., Wu, Z., Gordon, D., Eisen, J. A., Paulsen, I.,
 Karp, P., Komoro, F., Zhang, S., Yoo, H., Rao, Y., Biddle, P., Jung, M.,
 Katspan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
 Zhao, Z., Dolan, M., Tingey, S. V., Tomb, J., Gordon, M. P., Olson, M. V.,
 and Neester, E. W.
 Direct Submission
 Submitted (27-SEP-2001) Department of Microbiology, University of
 Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
 98195-7242, USA
 Location/Qualifiers
 1..10707

```

/organism="Agrobacterium tumefaciens str. C58"
/mol_type="genomic DNA"
/strain="C58"
/isolate="U. Washington"
/db_xref="taxon:176259"
complement(138..473)
/gene="n1d"
/ncbi_synonym="Atu3901"
complement(138..473)
CDS

```

/gene="nirD"
 /note="identified by sequence similarity; putative; ORF
 located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="nitrile reductase small subunit"
 /protein_id="AA144709.1"
 /db_xref="GI:17742340"
 /translation="MDSWIDIGDISIPLRGARCVKTPNGKIAVFETNAQVPAID
 HCPHKGGLSGIVGASVTCPLHNWVLSLTKALGADGEVRTIPVRNLGRLSLA
 LETIMIAAE"
 complement (479. .2926)
 /gene="nirB"
 /note="synonym: Actu3902"
 complement (479. .2926)
 /gene="nirB"
 /note="identified by sequence similarity; putative; ORF
 located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="nitrile reductase large subunit"
 /protein_id="AA144710.1"
 /db_xref="GI:17742341"
 /translation="MTQKLVIGNGMAPGRMLNLFETAPGLYDVTI FNAEPRVNDYDR
 IMLSVLSGSESYEDI VHNDEWYANNVTLHKAKATYRDRANKIVTSNGSLTASD
 RLVAITSGLPETIIPVPGHLPGLVATYRDLDTVQMLIAGKGAIAIYGKGLBA
 YGLKRGQDVTIVHMLPTIMEROLDRPAAVYLEKALNERGIDITTKNTGCIIEGEKV
 EGILEPDRITIKGMVVMVAVGIRPASGLAEAGIANGVGVDDGMMTPDSATIAIYV
 CAERHGKCYGVAPLPSASRYLADRLRCGSAEHTGISTNTKLKATGINTLSAGDPAGC
 DDIRESIYKRDYAGVYRLLIKENRIIGAVLYGSTADGSMFEDLMKSTDISARETL
 IFGOVYQGGSPDPMAAVVALPDAETCGCNGVCKGKITSATISKLTSLDVAARPK
 ASASGNTCGTGEQEDLTMLTLTGDTYNPAVOPMCKDTGLGHDDVRLTKAAGLTIIPV
 MOELMKWTSCCGACRPALANYLYVCDPDEADYDQSRYNERNVHANIQDGTYSVPV
 RMWGQVADLRAIADVDKFEIPLKVTGGQRIIDLGIKEKEDLPVAMDLGAGETI
 SGQAVAKLIPVTKCGEOKRCRGTOSTGLGFRIEKPMGSGMTPALKIAGVCCPN
 CAETCKDITGYICVDSGFETHPAAGAGIDIKGTIVGLYHTEDEALHLYALTYMRE
 QARYERYTKAKRIIGDEVARQIMDADRKAFYRFRVFSQKFAQVDPWSERSVSGD
 KHEPMAAIFGFSAAE"
 complement (2932. .3729)
 /gene="nirC"
 /note="synonym: Actu3903"
 complement (2932. .3729)
 /gene="nirC"
 /note="identified by sequence similarity; putative; ORF
 located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="ABC transporter, nucleotide binding/ATPase
 protein [nitrate]"
 /protein_id="AA144711.1"
 /db_xref="GI:17742342"
 /translation="MNAVYLKLDHDKYFDKGGSAEVLKGINLTIEKGEFVSYIGHSG
 CGKSMPLILIGLTKVAGAVLYLEMRVNEEGPRAVFPQNSHLPSTLYEENVALV
 AKVPNTTETIRERDWMRNLDLVOMARADKKESEISGGKORVGLPRLAALPEKLL
 ILDEFGALDLITRAHLDQAMVEIHAIRGNMTWITTDVBAVLIASRIYMTNGPRA
 HIGEVLDVPLARPPNRRIELASDKIYLLKCRNAVLFYTERRRFVDAAE"
 complement (3743. .4636)
 /gene="nirB"
 /note="synonym: Actu3904"
 complement (3743. .4636)
 /gene="nirB"
 /note="identified by sequence similarity; putative; ORF
 located using Blastx/Glimmer"
 /codon_start=1
 /transl_table=1
 /product="ABC transporter, membrane spanning protein
 [nitrate]"
 /protein_id="AA144712.1"
 /db_xref="GI:17742343"
 /translation="WSLARKFNEEPQOAPVQKANYVALSPKHLPRIDLRKMGAAAR
 NTIPLVALVAVLGWOLCGAPSSLPSPQVFEESYDILIVSFFNGSQDGLGLR
 VLVISERAVYAGFAAIIAGITIVAGIVCSWAMKGLDPIQVARTVPLMLPLSLAA
 FQDPSNPATFIYFTITSIMPVYINTVAVVRNIPQDYRVNAVAKRLINQLEPFFKIMLP
 SA

```

gene
APY1FTGLRIGIGLSMLAIVAEMLTGCGVIGFPIWDAMNSRLPDIIVALAVIGVWG
FMDXIVAAIGKXITRGAAN"
complement (4750. .6042)
/gene="nrtA"
/note="synonym: Atu3905"
complement (4750. .6042)
/gene="nrtA"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="ABC transporter, substrate binding protein
[interate]"
/protein_id="AAL44713.1"
/db_xref="GI:17742344"
/translacion="MKKIFSGTVSRRTILKTTATYALVTAVRTAPSPGAPATAEDEV
KAGKLGFIALTDAAPLIIAEKGLFAKHGPEVEVLKQSWGATRNLYVIGASNGID
GAHLLTPMYLMTGKVTQNNVPVMTIILRLNDLQSGISVAKETGOLDASLTK
AAPEKKADGKEIKAAMTFPGGTHDLNIRWLAAGGIDDDKDVSTIVPPQVAMK
VGNMDEVCGEPENBOLVNOGIGFTACTTGELMKGPBKALGKRAMVETKPNATKAL
LMAVMAQOCDEMANKEKSTILGRQNFNPVKDVLGKLNINIVGKRVLENTGL
QMKFWQDHASVPPHSDSWFTENIRWGFADPTDVKALVAKVNRREDIWRMAAKDGV
ADLPASTSRGKEFFPDGKVFDPENPSAYLESISIKAS"
complement (6423. .7718)
/gene="nrtC"
/note="synonym: Atu3906"
complement (6423. .7718)
/gene="nrtC"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="ABC transporter, nucleotide binding/ATPase
protein [interate]"
/protein_id="AAL44714.1"
/db_xref="GI:17742345"
/translacion="MAAPEKTAGSKGATOMDAPALVGSROKILRAGFIPLVNDASV
LTAARFGRADRGITLTVKQVSNANVADLAFOPLAHMLSPVSMGSGNP
SPTTPSPGRGNATITLSTRLPARKALTGSETRGALNNALTLVDDMRGEA
PTLGTTPFSSHNTEFRWLAAGGIHPHDVGLVVPPTSDALAAGLDGFCEGA
PNIVIAERGVGIVAAKQDLWPSAPEKVIQMRPEWASQOETVGRILTLAALASWC
DLADNDALSGALADPRYIGAPQSIIRRYLAGSPSIDSGNRVIEKPTFGDHANY
PROSGSLTVSQMIRWGOAELSETGVNALSARPDIVRAALGDGKAPEDADIRIEQO
DGDPRVDSGVDPDADIAGVNSFAVRISVPVSABSGEA"
complement (7754. .8341)
/gene="nast"
/note="synonym: Atu3907"
complement (7754. .8341)
/gene="nast"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="two component response regulator"
/protein_id="AAL44715.1"
/db_xref="GI:17742346"
/translacion="WTLRDLTILVIDENAIRASIIIEGREGAGYQRTVYIHEVNGVAR
TLETLPDVIATIDLENPNDMMEHLFQLRTVGRPLAMVEDSDTASIEAAVEAGVSA
VYVDGLKKERVPIIDMAVSFRNAPSLRDLADAKSALEBRKVIERAKGILMKRGL
SEEAFLALROGANMEKKISIEIAGSVTAAGLWM"
complement (8443. .9138)
/gene="Atu3908"
complement (8443. .9138)
/gene="Atu3908"
complement (8443. .9138)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, GntR family"
/protein_id="AAL44716.1"
/db_xref="GI:17742347"
/translacion="MLHCSERSASMTETIEGLPLADRLAKGIRTLISGELAPGOLS
EAAFSBRPVSRRNSLREAFRLTLKQGLLHEANRGVFAVAVPTMSVIDIYRVRNVEC

```

```

Query Match          9.6%; Score 133; DB 1; Length 10707;
Best Local Similarity 50.2%; Pred. No. 1.9e-15;
Matches 572; Conservative 0; Mismatches 495; Indels 72; Gaps 7;

87 CGATGTCGGGGCAATGTCGAGCTGTGAAAGAAATTGGAATTCGCTTCAAGCT 146
   |||||
5941 CGCTTTCGGGGCCACCGCAAGAACCGGAGGTGCAAACTCGTTTCATCCGCT 5882
   |||||
147 CACCGCATGGCGCGCTGGCGGCGCGCCGAAAGAGCTTTCGAGAGCGAGCGCT 206
   |||||
5881 CACCGACGCGCGCGCTCATATTCGCGCCGAGAGGACCTTTTCCAGAGATGACAT 5822
   |||||
207 GTTCGTGCAACTGGAAGCG---CAGGCCAACTGGAAGGTGTGATGATAGGTCGTGA 263
   |||||
5821 GCCTGAGTTCGAGGTTCTGAAACAGGCTCTTGCGGGGGCGAGCGCGACATCTCGTGT 5762
   |||||
264 TGCCCACTGACAGGCTCGACATGCTGCGCGCGCGCTTACGCCGACGCTTGCTT 323
   |||||
5761 CGGCGG-----CGCTTCAGCGGCATCGATGGCGGCATATCTGACCCGATGCCCTTA 5708
   |||||
324 CGGCAACGAGCGCGCATATGAGGTGCGCTTCAGCATGGGCTTCAACGCAAGCGCATTC 383
   |||||
5707 TCTCATGCACTACCGGCAAGGTGACGAGAACACGTCGGTGCATGATCCTCCG 5648
   |||||
384 GGTGTCCAATGAATTCGATCGATCGATGAAGCGCAATACGCTGGAAGCGGTAAAC 443
   |||||
5647 AGCGCTTAATCTCGACAGCGCAAGGCAAT-----TTCGCTGCGCAAGAAATATGCAAGAC 5594
   |||||
444 GGTGCACTCGATCAAGCGCATATCTTAACCCGCTGTCGAAAAATCAAAAGCCGAG 503
   |||||
5593 CGGCTGCAAGCTCGATGCTTCGCAAGGTGAAGCGGCTTCGAGAAAGAAAGGCGCACG 5534
   |||||
504 CAAGCGCTCAATATGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 563
   |||||
5533 CAAGGAATCAAGGCGCGCATATCTTCCCGGCGGACGATGATCTGATCTCGCTTA 5474
   |||||
564 CTGGCTGGCGGCTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
   |||||
5473 TTGGCTGGCTGCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 5440
   |||||
624 CCAATCGCGCAGACGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGAT 683
   |||||
5439 -----GACGTCTCCACCATGCTGCTTCCGCGCGCAATGCTGCAAAACAT 5393
   |||||
684 GGAAGCGCGCACATTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTT 743
   |||||
5392 GAAAGTCGCAACATGACGCTTCTGTGTAGCGAGCCGTGGAACGACAGCTCGTTAA 5333
   |||||
744 CAAGGCGATAGCGCTGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 803
   |||||
5332 TCAGGCGCATCGGCTTTATCTGCTGTACCAACCGGGAACCTTGGAAGAGGCAATCCGAAAA 5273
   |||||
804 AGTCTTGGCGGTGACCAAAACAATGCGCGGAGAAATACCCCAACCTTATCTGGCGGTAC 863
   |||||
5272 GCGGCTCGGCAAGCGCGCGCACTGGGTGGAAGAACCCCAACGCAAGGCTTTGCT 5213
   |||||
864 CAAGCGCGATGATGCGGCGCGCATGCTGCTGGAACCGCAATTAACAAGAACCGCAAGA 923
   |||||
5212 GATGCGCGGATGAGGCGGCAACAATGCTGCGCAAGATGCGCAACAAGAGAGATGTC 5153
   |||||
924 AGCCATCGAAATGCTGCGGCAAAACAATACCTCGGTGCTGACGTGGAAGTCTGGCGG 983
   |||||
5152 CACCATC-----CTCGGAAAGCGCAAGGTGTTCAACGTGCGCGGAGAGAGTCTCGG 5099
   |||||
984 TAGCATGAACGCACTTTCGAATACGAAGAAAGAGATTAACGCGGCTTACCGGACTT--- 1040
   |||||
5098 CCGCTCAAGGCAATATTAATTAATTAACGCGCGGCTGTAAGAAATATCTGCGCTGCA 5039
   |||||
1041 CAACACTTCTTTCGCGCGGCGGCAAGTATCGCTTACAGCACTGCAAGTGTGATCT 1100
   |||||

```


TANKDALDASSAARVAAWAKERGLSVGTNGCEDITLHAGHVSILHAASOCRLVIG
 LNSDASRLKGGKRPVNDHDBACIALALASPDANVVEEPRLIALILALPDLIVK
 GADYITDITVGDADVOKAGRVYLDVLVAGKSTGTITGLKRGSTS"
 complement (6734. .7366)
 /gene="m112564"
 /complement (6734. .7366)
 /gene="m112564"
 /note="hypothetical protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB49663.1"
 /db_xref="GI:14023056"
 /translation="MPSTELLAFPATTAIPAYIPGPAMLYAAOITLARGMSGLTA
 LCHIGGVHVPAALAGSLVPLHAPVPLMYAKYLGAALILMGSLFKRREGVVAL
 PALERSARAPFESITTEVLYNPKTILFMAFLPQDLSALPWLQFVGLTIVNL
 MFSADIVCFPLGLVLSRLRRSSRAQRMVQAGAVLGLGVALQRS"
 7499. .8449
 /gene="m1r2565"
 /gene="m1r2565"
 /gene="m1r2565"
 /codon_start=1
 /transl_table=11
 /product="ADP-L-glycero-D-mannohexose-6-epimerase"
 /protein_id="BAB49664.1"
 /db_xref="GI:14023057"
 /translation="MTIYVGAGMTGSNTVAALMAGHDILVVDLTGHTANLAD
 LSIADYLDQDFLTRIEAGALGRLEAVFQACSTTENNKFMEVNAVRYKRLHA
 COAPRVPLVASSASVYGGSEFEREPALERPLNVAYSKPLFDVVRVTDHSLA
 VAGLRFVNYGPREAHKGMASVAPHLFNOVERGENPRLFGADGPGQSRDFHV
 GVADVNLMLMKRGSSGIFNCGTGAQPRALAEVITDILGKEIIFEPDLKQSY
 OSFTQADMSRLAAGVNGOFRITETGVRYVEMLKQRS"
 8422. .9429
 /gene="m1r2566"
 /gene="m1r2566"
 /gene="m1r2566"
 /codon_start=1
 /transl_table=11
 /product="ADP-heptose-LPS heptosyltransferase"
 /protein_id="BAB49665.1"
 /db_xref="GI:14023058"
 /translation="MAESPILLIVIGPRWVDMVMAOCLFSALKELYPNALIDVLA
 AAPLVKMEPIRQIDPLPKPGALERTIRRGRLGRDVAAYVPGWKSALIPF
 ALIPRFGHLREMRVGLTDIYLPDAVRSYTAQVFSALGSSPAAPLITADNQA
 ALIDPRGLAPRFAALMPEAGFERGPAKWSSESAYAGAKRFMKGLKVALFGKNDQV
 THEITLAPGVVDLQOTRLEDAIDITAAALAVSDSLMVAAGVGPVIAVVGST
 SPENTPPLAERRELWLHLSCSPCHQKVCPLGLHNCCLKTLEVGAVATAERLLEPAT
 A"
 9426. .10400
 /gene="m1r2567"
 9426. .10400
 /gene="m1r2567"
 /note="unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB49666.1"
 /db_xref="GI:14023059"
 /translation="MKVLIVKTSMSGDVHTTPPAVEDASLHRPDISPDKCEVBAFAGI
 VALHPAIGIHTVALIRMRKALFNGGTWEEAALRRTRLECFEDVITAOGLKALV
 AAOARAPILGPDSSARBSATLFYVYVGTGWRNLAIERTRLFGALGYPDLSTL
 ASGIVPPGALAGITGKTAFLLHGTSSREDKMPVEDMIGTARLVERGTPPTWSNE
 REKSVAEIARAIVPVVVKSPADIALLGRSTVLGADTGLTHLSAFGLPIYAV
 PLATPREGLTGPRGPYASTILAVPGGRILTPAEVVAERDLSLNLKQAPA"
 complement (10535. .11008)
 /gene="m112568"
 /complement (10535. .11008)
 /gene="m112568"
 /codon_start=1
 Query Match 8.8%; Score 122.8; DB 1; Length 348077;
 Best Local Similarity 51.7%; Pred. No. 1.7e-13;
 Matches 394; Conservative 0; Mismatches 317; Indels 51; Gaps 3;
 455 TCAGGACATATATCAAAACCGGTCGTGGAATAATACAAAGCCGAGCAAGCGCTTCA 514

Db 227011 TCGATGCTCCAGACTGAAGACGCTTGCCGCCAAGAGGACCGGCAAGAGCTCA 227070
 Qy 515 ATATGGCGATGACCTTCCCGGCGGATCTCAACAATCAACTGCTTACTGCTGGCGG 574
 Db 227071 AAGCGCGCATGACCTTCCCGGCGGAGCAGATGACCTTGGGCTGCTACTGGCTGGCGG 227130
 Qy 575 CTGGCGGATCAATCTGCGCTACTATTTGGCCCGCGGAGACATTTCCGGCCAAATGGCG 634
 Db 227131 CCGGCGGATCATCTCCGCAAG----- 227153
 Qy 635 CAGAGCGCTTGTGTGGAGACCCCGCGCGGAAATGCGTCCAGCTGGAAGCGGCA 694
 Db 227154 --GAGCTTCGACCATCTGCTGTCGCCCGCGCATGCTGGCCATGAAAGTGGCGCA 227211
 Qy 695 CCATTTTGGTATATGCGTGGCGAGCCGTGAAACGAGAGCCGTGTTCAAGGGATAG 754
 Db 227212 ACATGAGCGTCTTTCGCGTGGCGAGCGGTGAAAGAGCAGCTGTCACAGGGCGTGC 227271
 Qy 755 GCGTGGCGGATCACTCCATGAAAGAACTTGAAGACAGCCCGGAAAAAGTCTTGGCG 814
 Db 227272 GCTTTACCGCCCGCACACGAGCGGATTTGTGAAGGCTCATCCGAAAAAGCGCTGCTC 227331
 Qy 815 TGACCAAACAATGGGCGGAGAAATACCCCAACCATATCGGCGGTGACCAAGCGCTGA 874
 Db 227332 TCGCGCGCGAATTCATCGAAGAACCCGACCCCAAGGCGATCTGATGGCCGTCA 227391
 Qy 875 TTCGGCGCGCATCTGCTGCGAGCGCGCAATACAGAAACCGCAAGAGCATCGAA 934
 Db 227392 TGGAGCCGACGAGTGTGCGAGGCGCAAGAGAACAGAGATGCGCGCATCATCG 227451
 Qy 935 TCGTGGCGCAAAACAATTCGTCGCTGCTGACGTGAAGTCTGGCGGCTAGCATGAACG 994
 Db 227452 GCAAGCGGCAATGATGAACGTCGCGCTGCGCAATCATTCGCGCCCTTAAAGGCGACA 227511
 Qy 995 GCACCTTGAATACGAAAAAGACGATTAACGGCGCTACCGGACTTCAACACTTCTTC 1054
 Db 227512 TCAACTACGCGAATGCCCGCGTGCACAGGCAAC--CGATTTTCAATGAAGTTCTGGA 227568
 Qy 1055 GCCACGCGCGCAGCTATCCGTCCTACAGAGTGAAGTGTATGACCCAGCTGAGCG 1114
 Db 227569 AAGGCGGCGGTTCTTCAACCCCTTCAAGAGCATGATGCTGTTCTCGCGCAAGCATCG 227628
 Qy 1115 GCTGGCGCATGATCAATGAATCAACCGGACATCTGATTCGATACCGCAAGACG 1174
 Db 227629 GCTGGGCGCA-----AATTCGCGCGCACGACGATCAAGCGCGTGTGACCAAG 227679
 Qy 1175 TCTACCGCGGACATCTATCTGCGCGCGGCGCAAGAACTGG 1216
 Db 227680 TCACCGCGAGGATCTGTGGCGGAGGCGGCGAAGGATCTCG 227721
 RESULT 10
 AP005945 300150 bp DNA linear BCT 12-SHP-2003
 LOCUS Bradyrhizobium japonicum USDA 110 DNA, complete genome, section
 DEFINITION 11/31.
 ACCESSION AP005945 BA000040
 VERSION AP005945.1 GI:27350985
 KEYWORDS
 SOURCE Bradyrhizobium japonicum USDA 110
 ORGANISM Bradyrhizobium japonicum USDA 110
 Bradyrhizobium japonicum USDA 110
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Bradyrhizobiaceae; Bradyrhizobium.
 REFERENCE
 AUTHORS Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uehi, T.,
 Saesamoto, S., Watanabe, A., Iidesawa, K., Itiguchi, M., Kawashima, K.,
 Kohara, M., Watsunoto, M., Shimpo, S., Tsunokawa, H., Wada, T., Yamada, M.
 and Tabata, S.
 TITLE Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 Bradyrhizobium japonicum USDA110
 JOURNAL DNA Res. 9 (6), 189-197 (2002)
 MEDLINE 22484998

PUBMED 12597275
REFERENCE 2
AUTHORS Kaneko, T., Nakamura, Y., Sato, S., Minamiawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iritaguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimo, S., Tsunooka, H., Wada, T., Yamada, M. and Tabata, S.
TITLE Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110 (supplement)
JOURNAL DNA Res. 9 (6), 225-256 (2002)
PUBMED 22485002
PUBMED 12597279
AUTHORS 3 (bases 1 to 300150)
JOURNAL Kaneko, T.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; 2-6-7 Kazusa-kametari, Kisarazu, Chiba 292-0812, Japan (E-mail:kaneko@kazusa.or.jp, URL:http://www.kazusa.or.jp/rhizobase/, Tel:81-438-52-3935 (ex.2338), Fax:81-438-52-3934)
FEATURES
source Location/Qualifiers
1.300150
/organism="Bradyrhizobium japonicum USDA 110"
/mol_type="genomic DNA"
/strain="USDA110"
/db_xref="taxon:224911"
/gene="b112729"
complement(14. .862)
/note="ORF ID:b112729
probable ABC transporter permease protein"
/codon_start=1
/transl_table=11
/protein_id="BAC47994.1"
/db_xref="GI:27350986"
/translation="MDNRERMTPTROIIGKIGLMSFVIVSPALFELMMLSLK
FEIDNAAYPVFEPEHIAWKYADVLASNRPLTYFNSLVTSATLALIVGPDGY
GIARMAHKSATVILARITPGISYLIPLFLPOMLGLGLTVPOIILHVVTVIVI
WMIGCFETTPLELEBALIDGATRWQFRHVALPAKGIAPAFILAVIFSMNPFV
GIVLAGRETRTIPVAVYTMISFDQSLMGPLAALIVTLPULLTVFAQRQIVAGLTA
GAVKGG"
complement(835. .1773)
/gene="b112730"
complement(835. .1773)
/gene="b112730"
/note="ORF ID:b112730
probable ABC transporter permease protein"
/codon_start=1
/transl_table=11
/protein_id="BAC47995.1"
/db_xref="GI:27350987"
/translation="MSALTQADPAAARSDAPEKELRPSYFPVVVPALVVVLAITIF
PWFPTIMSLNWKGSPTTFVGLSYLRLTSDPRTEVGHRTVYTALSVLPLVLG
TAAAVFHQKFAQGRGLRGVFIIMPMAATPVALATWIMVHPOQLGVNLTSLVGI
QLVAVHAFPIPSLVLEVETQWTPLVMLVLGSLAIPIEPESAOIDANIMQMERF
ITLPLIMPLFIAGMIRMTDAVKSFDIIRAIIOGGGSGSETINITYLSVAVPTYYDLG
VGSATVAVPELLIVLAAVMLYARQRMWVTEIASDA"
complement(1770. .3152)
/gene="b112731"
complement(1770. .3152)
/note="ORF ID:b112731
probable ABC transporter substrate-binding protein"
/codon_start=1
/transl_table=11
/protein_id="BAC47996.1"
/db_xref="GI:27350988"
/translation="MARSLVSLPKTMSHREDPEVVDYRSRTLLTGTAAGALSITG
LBARAVNMKKYAGTKLEVLAKPGRGDLAKNIKEFELNGIOVEABQIPROOROK
VVIETSGRPSFDVVAHLSTHYOKROPEKAGMLADTTPFKDPSLTAPDLVESDFNAG
LDYAKNDKQMLSLPWSVDYFTLYNKELFQKRGAVAPKTLDEMVAAREKLTDPKEGT
YGFVGRGLNNAMLMTNFNLVYGGFLLDDKNNIITLDGEALIAATKLYOTLLTKSAPP

gene
CDS
complement(3179. .4012)
/gene="b112732"
complement(3179. .4012)
/gene="b112732"
/note="ORF ID:b112732
putative cytochrome C"
/codon_start=1
/transl_table=11
/protein_id="BAC47997.1"
/db_xref="GI:27350989"
/translation="MDRRORL.FVHRLGITDHGRMKFWRALAVATLLAAAPALLAGRIAP
BARBSGYAFMGEDTRAMQDDDTANBGM.FYLDGELMFKTKGSARAKACADCGDARG
MKGVAAYPAREKAVGRPTITLDQRIINLCRANHQQATPLPYESRDILLASFAVAHDSRG
VAITAGDDPOLKQFVQGRDLFMQREGLNACTCHDNDPDRHAGAPITQAOPTGY
PLYRLMEQTLSLERRLRSCMSGVAQAYDGAPELVALLLEYILMSRARGLPMETPAVR
P"
complement(3955. .4272)
/gene="b112733"
complement(3955. .4272)
/gene="b112733"
/note="ORF ID:b112733
probable sulfur oxidation protein"
/codon_start=1
/transl_table=11
/protein_id="BAC47998.1"
/db_xref="GI:27350990"
/translation="MAALINVPAKARQGVIEIRLTLSHINMETGRHATADGLVPRDI
ITSFTCRYNGREVFRAIDLPAIAANPYLSFTVAKESGFEFEMTGDNYSSTASASI
TVE"
complement(4272. .4739)
/gene="b112734"
complement(4272. .4739)
/gene="b112734"
/note="ORF ID:b112734
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC47999.1"
/db_xref="GI:27350991"
/translation="MPTTRQFLSLAGSITAGTIPITLRLPLEATPAMLNTAIRNV
GEAPIRTRGVKLDIPLVENGNTVPMSTSVASPMPTDDYVSGSHVENKNPQNPNGF
HLGPWAGSAQVSTIRILADTKVVALIALSDSSFMQVADVVYTLTAETERN"
complement(5385. .6665)
/gene="b112735"
complement(5385. .6665)
/note="ORF ID:b112735"
/codon_start=1
/transl_table=11
/product="flavocytochrome C flavoprotein subunit"
/protein_id="BAC48000.1"
/db_xref="GI:27350992"
/translation="MRPMNAVPTRRNVVIGITTAATSYAVPSIILBAQSAGRVVVVGG
FGGAACALTRANKAQVTLIEPKVGVISCPFSNEVYAGIRETEAQQRGYDKYVAG
ITVVAQAATTYDPERRSVTTTDDGATLITDRLVLSGIDPFHYBALPGYDEASERKPHA
WKGAQOTLLRRQLEAMEDGTVALAIAPANSRCPAPAYEASLTAHLTKTKPPAKV
LILDAKDSFSOQRLFERAMKELYGMIMERIGLSQGRVTSVASTKTIIVTERGNTPD
VANVPIPORAGHLAQIAGADATGWCPIDPVTFESKLVPNHIVIGDACIGGCI PKGS
AASQGRCAAAIVILAGRAPETPRFLGVCTTVAAPGYGSLAGNYQPRGDIYAWE
GGATSPVDAPREIARREADEARMFQTTADTFG"
complement(6782. .8968)
/gene="b112736"
complement(6782. .8968)
/note="ORF ID:b112736
putative aldehyde dehydrogenase protein"
/codon_start=1
/transl_table=11
/protein_id="BAC48001.1"

```

/db xref="GI:27350993"
/translation="NMKHVSPKMRRAFYIATVAGAGLIGLIDIPROGPAYVRADG
SEBISAMVYVRPDUTYVIRIASBEMQGSILTELAQIVAELECDMSKSTERYPTGQS
VAKRVMGDFSTGSGRISQDYVRKGATVRVMIQKAAOSMKVPASECTAAASVI
THTPSGRITTYGVKVAEAAKLTPEADVTKADPRMVLIGKVLRLDTADKTVMWY
IDVKLPGMLNAAIKDCPVFGKLSYDEAKIAGMKGVKVVLTGTAVAVVADTWWHA
KPALEALPIVMEGDGAKVSSASIAKMLBGLDNQGPAPVAVKNGDAPAKAIAGVKY
EAVVYVPIVYNHATMEPMNATALTAEKCEVNGCTONGEAPAAITLEAGLPAEKDVH
KMPGGGFRRCQGTQYVROAVIAKOMPTPIKLMSREDAHGRHYHTTQCKRTGA
PADNNVVALHRLSGOSTLFSLRPEALQNDPAPAFQVAGSGEAAGFYVSNLLIE
HAMRNPHPVPGFWRGVNVMNAIYMECFMDELAAAGDPLEFRRLKSNHFKLAVL
NAVAPKIGWTTAPQGVYRGIQVWVGSGYVAGAAEISYDGNKTLKVHIVASTPGY
VNPAPVERIOIAGSFYVGSALFYGGCTYKDGKIDQNTDPTNSMRINEMPVESVMV
BSGGFMGVGEPTTGYAAPAVIYNAVFAAIGKIRIRVPLRDQNTTA"
/gene="b112737"
/complement(8980..9456)
/gene="b112737"
/complement(8980..9456)
/gene="b112737"
/translation="MDMTQGTAMNLTJNGKSLTLDVEPDPTPLMAIRENGLTGT
VCGGIACGCACTVHIDGATRAQGVSAABEKKITTEGLASGSLHRYQAMTAKD
VQCGICQSGMTMAVVAALINERKPTDADIDAITNICRGTFQVREAIHTTASA"
/complement(9678..10070)
/gene="b112738"
/complement(9678..10070)
/gene="b112738"
/complement(9678..10070)
/translation="ORF ID: b112738
unknown protein"
/codon_start=1
/translation=11
/product="oxidoreductase with iron-sulfur subunit"
/protein_id="BAC48002.1"
/db xref="GI:27350994"
/translation="MDMTQGTAMNLTJNGKSLTLDVEPDPTPLMAIRENGLTGT
VCGGIACGCACTVHIDGATRAQGVSAABEKKITTEGLASGSLHRYQAMTAKD
VQCGICQSGMTMAVVAALINERKPTDADIDAITNICRGTFQVREAIHTTASA"
/complement(9678..10070)
/gene="b112738"
/complement(9678..10070)
/translation="ORF ID: b112738
unknown protein"
/codon_start=1
/translation=11

Query Match      8.6%; Score 119.8; DB 1; Length 300150;
Best Local Similarity 52.4%; Pred. No. 6.4e-13;
Matches 404; Conservative 0; Mismatches 307; Indels 60; Gaps 4;

470 TAAACCCGGTCGTGMAAAATACAAAGCCGAAGCCAGCCGTTCAATATAGCGATGACCT 529
80688 TGAAGCGCGGCTTCGAGAAAGAAAGCCGAAGCCAGCCGTTCAATATAGCGATGACCT 80747
QY
530 TCCCGCGCGGATCTCAACAATCAACTGCTTACTGCTGCGCGCTGCGGATATCAATC 589
80748 TCCCGCGCGGATCTCAACAATCAACTGCTTACTGCTGCGCGCTGCGGATATCAATC 80807
QY
590 CTGGCTACTATTGCGCGCGGAGACATTTCGGGCCCAATTCGGCGGACAGCGCTTGTGT 649
80808 CTGACAG-----GACGTGAGACCA 80828
QY
650 CCGTACCCCGCGCGCAAAATGCGTCCACGCTGGAAGCCGCGCAACATTTTCGTTATT 709
80828 TCAAGCTGCGCGCGCGGAGATGTCGCAACATGAAGTCCGCAACATGAGACGCTTCT 80888
QY
710 GCGTGGCGAGCCGTGGAACACAGCAGCCGCTGTTCAAGGGCAATAGCGCTGCGGTATCA 769
80888 GCGTGGTAGCCCTGGCCCGGAGCTGTGTCACACGAGGCATTCGCTACAGCGCGCTCA 80948
QY
770 CCGATGAAGACTCTGGAAGAGACGCGCGGAAAAAGTCTTCGCGCGTGAACAAACATGGG 829
80948 ACAACGCGAGATCTGGAAGAGACATCCGAAAAAGTCACTCGGCAATCGCGCGCGCTGGG 81008
QY
830 CCGAAGAAATACCCCAACACTTATCTGCGGTGACCAAGCGCGCTGATTCGCGCGCGCATC 889
81008 TCGAACAAGATTCGAAGAGCGCGCGCATCTCTGATGCGATTAATGAGAACCGACAT 81068
QY
890 GCGTGAAGCCCGCAATAACAGAACCGCAAGAGACCATCGAATGCTGCGCGCAAAAC 949
81068 GG-----GCCGACAAGATGAGAACAGAGAGAGTCCGCAACCATCATGAGCGCAAGCGCG 81122
QY
950 AATAGCTGCTGTGACGTGAAGTGTCTGCGCGCTAGCATGAACGCACTTTCGAATAGC 1009

```

```

DB 81122 AATGATATCAATCGCCCGCGCGAGCATGCGCGCATGCGCCCAAGGCAAGTTCGACTAG 81182
QY 1010 AAAAAGCATTAACGCG-----CGTACCGGACTTCAACACTTTTTCGCCACGCGG 1063
DB 81183 GCAAGCGCGGCAAGGTGTGCGAAGATCGCCGCAATCATGAAGTACTGCGCATCAG 81242
QY 1064 CAGCTATCCGCTCTACAGCATGATGCAAGTGTGATCTGACCCCAAGAGCTGTGGGGA 1123
DB 81243 CCTCTATCCGTTCCAGAGCCATGATCTCTGTTGATGACCGAGACATCCGCTGGGGA 81302
QY 1124 TGATCAATGAATTCAACCGGACCACTGGTATCTGATACCCGCAAGAGCTGTACCGCG 1183
DB 81303 -----AATACGAGCGCGCTTTCGACAGAGCGCGTGTATCCGCAAGTTCACCGCG 81353
QY 1184 CCGACATCTATCTGCGCGCGCGCAAGAACTGTGCGAGAGGCAAGGCCA 1234
DB 81354 AGGACCTCTGGAAGAGCGCGCGCAAGAGCGCTTGTGTGCGAAGTTCGCGCA 81404

RESULT 11
RME591985_07/c
WPCOMMENT
Sequence split into 17 fragments LOCUS RME591985 Accession AL591985
Fragment Name Begin End
RME591985_00 1 110000
RME591985_01 100001 210000
RME591985_02 200001 310000
RME591985_03 300001 410000
RME591985_04 400001 510000
RME591985_05 500001 610000
RME591985_06 600001 710000
RME591985_07 700001 810000
RME591985_08 800001 910000
RME591985_09 900001 1010000
RME591985_10 1000001 1110000
RME591985_11 1100001 1210000
RME591985_12 1200001 1310000
RME591985_13 1300001 1410000
RME591985_14 1400001 1510000
RME591985_15 1500001 1610000
RME591985_16 1600001 1683333
Continuation (8 of 17) of RME591985 from base 700001 (AL591985 Simorhizobium meliloti 10

Query Match      8.5%; Score 118.2; DB 1; Length 110000;
Best Local Similarity 48.7%; Pred. No. 1.3e-12;
Matches 487; Conservative 0; Mismatches 463; Indels 51; Gaps 4;

132 CCGCTTCATCAAGCTCACCGAGATGCGCGCGCTGCGCGCGCGGCAAAAGCTTCT 191
56467 CCGATTCAATCCCTCTCGTGAATGCGCGCGCTGATTCGCGCGCGGCAATTCGCGCTTGC 56408
QY
192 CGAGACGAGGCGCTGTTGTCGCAACTGGAACGAGCGCAACTGGAAGTGTATGGA 251
56407 GGAAGCGGAGGATCACCTCTGAACTGTCGAAGACGCTCTCTGCGGCAATGTGCGCA 56348
QY
252 TAGGCTCGTAATGCGAACTGAGACGCTCGCAATGCTGCGCGCGCGCTTACGCGC 311
56347 CAGGCTTCGCTTTCGTCATGTTGAGCGTGGCCACATGCTTCAACCATGCGCTTGC 56288
QY
312 CAGCGTTGCTTTCGCGCAACA---GGCGGATTCGAGTGGCGCTTACAGATGGCTTCA 368
56287 CATGCTCGGCTCGGCTCCAAATCCGTCGCGCGCATGTCGCGCTTTCCTGAGGCGCG 56228
QY
369 CCGCAACCGATTAAGTGTCCATGAATGAATGATGATGATGAGAGCGCAATACCGCT 428
56227 CCGCAACCGATTAAGTGTCTGCAACGCGCTCTAGAGGCTCAAGAGCGGAGGCT 56168
QY
429 GGAAGCGGTAAACCGGTGATCGATCAAGCGAGATTAATTAACCGGTGCGGAAAA 488
56167 CCGTGAAGCGGAGATGCTGGAAGAGCCCAAGCGCGCTGCGCGCGCTATCCGAGCGC 56108
QY
489 AATCAAGCGGAGGCAAGCGCTTCAATATGCGCATGACCTTTCGCGCGGATTCACAA 548

```

```

Db      56107 GGGGGAGCGCGGAGAGCCGCGCTGACGCTCGGCGCTACACTACCTTTCGTCGCCATTA 56048
        |||||
Qy      549 CATCAACTGCGTTACTGCTGCGGCTGCGGATGCAATCTGCTACTATTGCGCGCC 608
        |||||
Db      56047 CTACAGGTTCCGCTACTGCTGCTGCGCGCGCGCATCTGATTC----- 56006
        |||||
Qy      609 GAGAGCATTTCCGGCAATGCGGCGCAGCGCTTGTGTGCGGTGACCCCGCGCGCA 668
        |||||
Db      56005 -----CGACCGCACTGCAAGCTGCTGTGTGCGCGCAAT 55967
        |||||
Qy      669 AATGCCGTCCAGCTGGAAGCCGCGACATTTTCGTTATTGCTGCGGAGCCGTGGA 728
        |||||
Db      55966 GACCTCGAGCGGCTGCGCGCGCGCGCATTCAGCGGCTTTTGTGCTGCGGCGCGTGA 55907
        |||||
Qy      729 CCAAGAGCGCGTGTCAAGGCGCATAGCGTCCGCGGTATCCGATGAAGAACTTGGA 788
        |||||
Db      55906 CATGTGCGATCCGAAAGCGCGCTCGCGCATCTGCGGACCAAGACATCTGCGC 55847
        |||||
Qy      789 GGAACAGCGGAAAGTCTTGCGGCGTGAACAACAAATGGGCGGAGAAATACCCAGAC 848
        |||||
Db      55846 CTGCGACCGGAAAGTCACTCGCATGCGTCCGATGAGCGAGGCTAATAGGATAC 55787
        |||||
Qy      849 CTATCTGCGGTGACCAAGCGCTGATTGCGGCGGATCTGCTGAGCGCGACATTA 908
        |||||
Db      55786 GGTCTGCGCGCTGTGCTGCGCGCTGACCGCGCGCGCTGTGCGAGCAACCGGCCAA 55727
        |||||
Qy      909 CAAGAAGCGGAAAGCAATCTGCTGCGGCAAAACAATCTGCTGCTGACGT 968
        |||||
Db      55726 CCGCGCGCGCTTG-----CGAAGTCTGGAAGAGCGTCACTGCGAGCGCGCGC 55673
        |||||
Qy      969 GGAAGTCTGCGGCTAGCATGAACGCACTTTCGATTCGAATAAGACGATTAACGCGC 1028
        |||||
Db      55672 CGATTCATCCGCGGCTACTTCCGCGCAATTCGAAATCGACCCCGAGGAAACGCGC 55613
        |||||
Qy      1029 GCTACCGGACTTCAACACCTTTCTTTCGCCACGCGC---GCCAGTATCTGCTACAGC 1085
        |||||
Db      55612 CGTCGCGAAGACTATCTGCTTCCACGAGGCTTTGCAACTACCGCGTCCAGCA 55553
        |||||
Qy      1086 TGCAGTCTGTATCTGACCCAGCTGAGCGGCTGCGGCGATGA 1126
        |||||
Db      55552 GCGGCTCTGACCTACAGCAGATGCTGCGCGCGCAGCA 55512
        |||||

RESULT 12
LOCUS   AL646058/c 196050 bp DNA linear BCT 02-SEP-2002
DEFINITION Ralstonia solanacearum GM11000 chromosome, complete sequence;
ACCESSION AL646058 AL646052
VERSION   AL646058.1 GI:17427202
KEYWORDS
SOURCE   Ralstonia solanacearum
          Ralstonia solanacearum
          Burkholderiaceae; Ralstonia.
REFERENCE
AUTHORS   1 Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
          Arlet,M., Billault,A., Brotier,P., Camus,J.C., Cattolico,L.,
          Chandler,M., Choisme,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
          Gaspin,C., Laveie,M., Moisan,A., Robert,C., Saurin,W., Schex,T.,
          Sigler,P., Thebaud,P., Whalen,M., Wincker,P., Levy,M.,
          Weissbach,J. and Boucher,C.A.
          Genome sequence of the plant pathogen Ralstonia solanacearum
          Nature 415 (6871), 497-502 (2002)
TITLE     Genome sequence of the plant pathogen Ralstonia solanacearum
JOURNAL   NATURE
MEDLINE   21681879
PUBMED    11823852
REFERENCE 2 (bases 1 to 196050)
AUTHORS   Boucher,C.A.
TITLE     Direct Submision
JOURNAL   Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
          Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
          Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,

```

```

COMMENT   BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
          Dausset-CEPH 27 rue Juliette Dodu, 75010 Paris, France, IMGM CNRS
          118 Route de Nardonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
          URGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
          Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
          F31326 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire
          INRA, BP27, 31326 Castanet-Tolosan Cedex
          Christian Boucher@toulouse.inra.fr
          http://sequence.toulouse.inra.fr/R.solanacearum.html.
FEATURES
          source
            1..196050
            /organism="Ralstonia solanacearum"
            /mol_type="Genomic DNA"
            /strain="GM11000"
            /db_xref="taxon:305"
            117..1154
            /gene="RSC0194"
            /note="synonym: RS00626"
            117..1154
            /gene="RSC0194"
            /EC_number="1.1.1.-"
            /function="small molecule metabolism"
            /note="Product confidence : probable
          Gene name confidence : hypothetical
          predicted by Codon usage
          predicted by Homology
          predicted by Framed"
            /codon_start=1
            /evidence=not_experimental
            /transl_table=11
            /product="PROBABLE ZINC-DEPENDENT ALCOHOL DEHYDROGENASE
          OXIDOREDUCTASE PROTEIN"
            /protein_id="CAD13722.1"
            /db_xref="GI:17427203"
            /db_xref="GI:17427203"
            /db_xref="GOA:O8Y2Y7"
            /db_xref="UniProt/TREMBL:O8Y2Y7"
            /translation="MKALVTEGPKISLENPKPELOAPGAVVVTLLTTCGTDLHI
          IKGDPVCEPRTGHEGVIESGATSLTSPKPDHYLISCSGCANCRGMY
          HCEHGWTLGRIDGVOAEYVRVPAETSLKVPQVDEALNMLSDIPNCEFGV
          NGRVPGSTVAIVGAPITGLASLTLPAQFSPRIITMIDTVARLEVASGLGSDCIN
          TRENAAAVVKRLTDMGVDCALIVAGVATPELCEQVAPGVANIGHGVKADLH
          ETLMSONAIITRLVDIVSTPMLKTVAGKLLPQOLITTHRFKLDIILAAVDTFNNAG
          ETHALKVIIT"
            1164..3017
            /gene="RSC0195"
            /note="synonym: RS00627"
            1164..3017
            /gene="RSC0195"
            /function="miscellaneous; hypothetical/global homology"
            /note="Product confidence : hypothetical
          Gene name confidence : hypothetical
          predicted by Codon usage
          predicted by Homology
          predicted by Framed"
            /codon_start=1
            /evidence=not_experimental
            /transl_table=11
            /product="CONSERVED HYPOTHETICAL PROTEIN"
            /protein_id="CAD13723.1"
            /db_xref="GI:17427204"
            /db_xref="GOA:O8Y2Y6"
            /db_xref="UniProt/TREMBL:O8Y2Y6"
            /translation="MRRRLRGTPGMMQAGMLTREPESLSPVCAVILADRYLARGETTP
          EGVRRVVAHLLAAAEAPDRTRHPRLFTVNLHGAICGRIMANAGVROATMNCV
          HPITVPGAQVAGADIDALKQAOVTLAGGAGVDFSVAPMDLQGEODGTSVC
          EIVIRFDTCAALALANTRGQAQAVLRCNHPDLAAFAARGRRKWTTFNVSAVTD
          AFQAVADDAQMRLLQRALEPSAQRRAAGHLIDNGHMCYATVPARQMDMLVEAHS
          AEPGLFTIDTINANDLAELFTIATINCGSPAPGSCVLPINLSMDIOPPGVS
          GPMEDPVSFGRVHIOVRM.DNALIDITRMRLIEMREAHAKRRTGCVGTGKADLTM
          MGLPQAPBARNLAVQIRCI.RDNAAYASASALAQGRSPPLFRARSLAPRGFVATLP
          GAVERTTARHGLRNHSLSLPTGSVSLAFGNCSSGIEPVDNAVYQRRVRIHQCPQ
          LYRIENAHYRFOAMHGHAVALPSYFVANOVDGDHDKMVAALQPTDASISKTLII
          SGVSPAVGALLFRWQLRLKGIITIFRPDPSDBEVLVLPADGRPVCFC"

```

```

gene
3343..4683
/gene="RSC0196"
/note="synonym: RS00628"
3343..4683
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE CYTOCHROME C1 SIGNAL PEPTIDE PROTEIN"
/protein_id="CAD13724.1"
/db_xref="GI:17427205"
/db_xref="GOA:O8Y2Y5"
/db_xref="UniProt/TREMBL:O8Y2Y5"
/translation="MSHLIALAKAFYMLLIGLLANSAGAVPAFAQOTGACMACH
VSEPLTPGRFPMKGYTLNSYTI.PVSGMVSTNRTTIDODFEDVRNGDTVPQ
QASLEFYAGRIVGCVGAQMTYDGI.AHSHALDNTDIRAAMHLREDDVDFYGTIVNN
FTVSDVMSPPAGYPPYAA.PNIGVPLAGLIDGAAQOVGASAVAFORHYVLAIG
AYRTADASFVPAQODTAPRGVABLAGANPYKAVYANQEGASHMTGTGILLANR
YEDNTMPTGTDRFDYAI.DAQYITMAHAFQVAMTYEHODMPSFSGIGAGP
TPANPTDLNAEFKAKASYVQRKYGATLAYFSTTGADPLYAQAAPVTSANGRPDTQ
GVIFELDYLPHPDQIKLALQYTWYLFKENGANVNDGNAINNTIYLLGMWFMF"
4719..5339
/gene="RSC0197"
/note="synonym: RS00629"
4719..5339
/gene="RSC0197"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE CYTOCHROME C SIGNAL PEPTIDE PROTEIN"
/protein_id="CAD13725.1"
/db_xref="GI:17427206"
/db_xref="GOA:O8Y2Y4"
/db_xref="UniProt/TREMBL:O8Y2Y4"
/translation="MKGFHQPRRVMLAGVITLATTGMSYTVAGNDSKLAALQACSH
GYHGSBSPFRDLQMPHYLEADLKGRDRGRTDQAQFMWGIASQUDVTTLASL
ABYVAKQTPAGAGADPALTAGCKE.FEGLPAGEPACASCHGAQAOQNDTFPRLAG
QHEAYLRQISVFKNCTRANAPVMSAVATLTDEQAKAVAFLOAO"
complement(5374..6015)
/gene="RSC0198"
/note="synonym: RS00630"
complement(5374..6015)
/gene="RSC0198"
/function="miscellaneous; unknown"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon_usage"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD13726.1"
/db_xref="GI:17427207"
/db_xref="UniProt/TREMBL:O8Y2Y3"
/translation="WVNLGLFVTVLNVIPAFAPPTMWNVSWIGTOPANPALLA
GIAACASTAGRLVLRSLCTLVRNMRMRADRONIDVIRTMWEGRSKMTVGVLVYAF
SFPNSNYFIAVGLTGMRLMLIGVPEFVGWRSYLTWTNIAQISARVYDESEINAGAY
MSIVFVLSQMLFALVYGPAKVDWHNIITKRLRMRRRRPPAPSSDDPRATND"
complement(6044..6568)

```

```

CDS
/gene="RSC0199"
/note="synonym: RS00631"
complement(6044..6568)
/gene="RSC0199"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Homology"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE CYTOCHROME B561 CYTOCHROME
TRANSMEMBRANE PROTEIN"
/protein_id="CAD13727.1"
/db_xref="GI:17427208"
/db_xref="GOA:O8Y2Y2"
/db_xref="UniProt/TREMBL:O8Y2Y2"
/translation="NRFEPKYSLGKRALHMLVFLAALIAVVTIELDAPPKSARAA
LFAAQITGLSVLALMLVRLVRVGQVPAPVSGSQLERRARLTHGVLYIMVGMPI
LGVLAIVMSGKBIQPFGLATLTLLADDTLANLAKVHESGATLVYIFVGLAAALW
HFIILKRLLRMI"
7807..8229
/gene="RSC0200"
/note="synonym: RS00632"
7807..8229
/gene="RSC0200"
Query Match
Best Local Similarity 54.5%; Pred. No. 3,4e-12; Length 196050;
Matches 305; Conservative 0; Mismatches 240; Indels 15; Gaps 3;
gene
637 GAGCCTTGTTCGTGATGACCCCGCGCGCAATGCCGTGACGCTGAGACCGGACCC 696
195304 GAGCGCAAGGCGATCACCGGTGCGCGCGCGAGATGTCGCAACATGGCGGTGGCAAC 195245
697 ATTTTCGGTTATTGGTCGCGCGAGCGCGTGAACCAAGCGCGGTTCGAAGGCGATAGCC 756
195244 ATGGAAGGCTACTGCGGTGGCGCAAGCGCGCGCGCGCCATGCGCGACGCGATCGGC 195185
757 GTCGCGGTATCATCCGATGAAGAACTCTGGAAGAGACAGCGCGGAAAAGTCTTCGCGCGT 816
195184 TTTCACCGCGAGACACACGAGCGCATCTGGAAGAACATCCGGAAGAGTGTCTCGACACC 195125
817 ACCAAACATATGGCGCGAGAAATATCCCAACCTATCTGGCGGTGACCAAGCGCGTGAAT 876
195124 AGCGCGAGTTTCGCCAAGAAATATCGAAACCGCGCGCGCGCTACACGCGCGCGTCTG 195065
877 CGGCGCGGATTTGCTGGAACCGGCAATTAACAAACCGCAAGAAAGCCATGAATG 936
195064 GAGGCGCTCAAGTTATCGATGCTCTCGGCTTGAACCGGCGCAAGACGAGC--CGAGAG 195008
937 CTGGGCGAAAACAAATAGCTGCGGTCTGACGCGGAAGTCTGGCGGCTTAGCAATGAACG 996
195007 GTGGCGCGCAAGTCTTCAACACCGACATGACATCATCTGGAACCGACATGCTCGGCG 194948
997 ACTTTCGAATACGAAAAAGACGATTAACGCGGCTTACCGGACTTCAACACCTTCTTTCGC 1056
194947 CGCTACACCAACGCGCTGGGCAAGACCTGGGAGCATGCGGACCGATGGCGTTCACAC 194888
1057 CAGCGCGC---CAGCTATCCGTCCTACAGACAGTGAAGTCTGATCTGACCAAGCTTAGG 1113
194887 GACGCGCGGCTCACTTCCCGTACCTGTCGAGCGGACATGTGTTCTTCAACACAGACAG 194828
1114 CGCTGGGCGATGATCAATGAATTCAACCGGACACCTGTATCTGGAATACCGGCGCAAGAC 1173
194827 CGCTGGGCGCTG-----CTGAAGACGACCCCGACCTACCTGCGCATGCGCAGGAG 194777
1174 GTCTACCGCGCGGACATCTA 1193
194776 GTCAACCGCGCTGACATCTA 194757

```

gene

RESULT 13	
LOCUS	AP005951/c
DEFINITION	298700 bp DNA linear BCT 12-SRP-2003
ACCESSION	AP005951
VERSION	17/11
KEYWORDS	Bradyrhizobium japonicum USDA 110 DNA, complete genome, section
SOURCE	AP005951.1 GI:27352606
ORGANISM	Bradyrhizobium japonicum USDA 110
REFERENCE	Bradyrhizobium japonicum USDA 110
AUTHORS	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
TITLE	1
JOURNAL	Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriuguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpou, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
MEDLINE	Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110
PUBMED	DNA Res. 9 (6), 189-197 (2002)
REFERENCE	2
AUTHORS	Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriuguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpou, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.
TITLE	Complete genomic sequence of nitrogen-fixing symbiotic bacterium Bradyrhizobium japonicum USDA110 (supplement)
JOURNAL	DNA Res. 9 (6), 225-256 (2002)
MEDLINE	22485002
PUBMED	12597279
REFERENCE	3 (bases 1 to 298700)
AUTHORS	Kaneko, T.
TITLE	Direct Submission
JOURNAL	Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research Institute, The First Laboratory for Plant Gene Research; 2-6-7
MEDLINE	Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
PUBMED	(E-mail: kaneko@kazusa.or.jp.
REFERENCE	URL: http://www.kazusa.or.jp/rhizobase/,
AUTHORS	Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)
TITLE	Location/Qualifiers
JOURNAL	1. .298700
MEDLINE	/organism="Bradyrhizobium japonicum USDA 110"
PUBMED	/mol_type="genomic DNA"
REFERENCE	/strain="USDA110"
AUTHORS	/db_xref="taxon:224911"
TITLE	48. .512
JOURNAL	/gene="b1r4344"
MEDLINE	48. .512
PUBMED	/gene="b1r4344"
REFERENCE	/note="ORF_ID:b1r4344
AUTHORS	unknown protein"
TITLE	/codon_start=1
JOURNAL	/transl_table=1
MEDLINE	/protein_id="BAC49609.1"
PUBMED	/db_xref="GI:27352607"
REFERENCE	/translation="MTQATPASTPKRRBPQWMLTDAAGRISETLQPADSETVG
AUTHORS	LKVGKNGCAGCAGSTVEYAHDPPTDEVEDGVKITLDPKAVFLILGRGGLGGRQ
TITLE	DAGPVRLQQPQPELRLRLRRIGRAAGGEDRRVGPALPFSAVVLAARHTYA"
JOURNAL	584. .1069
MEDLINE	/gene="b1r4345"
PUBMED	584. .1069
REFERENCE	/gene="b1r4345"
AUTHORS	/note="ORF_ID:b1r4345
TITLE	hypothetical protein"
JOURNAL	/codon_start=1
MEDLINE	/transl_table=1
PUBMED	/protein_id="BAC49610.1"
REFERENCE	/db_xref="GI:27352608"
AUTHORS	/translation="MGPGFRQDDDERGAMDBEFLIDFADPGPVTIRMGSGYGISA
TITLE	DGTFFALIRAGLFLRADELTIIDFETBSGKPFQYSTRAKTYVNSYWEILPARLPDSS
JOURNAL	OEIAQWRAALAAORAKVKORPKGKKASKVAFPKKVPARPKPAKKAARRVQKAKRAG

gene	R"
CDS	complement(1134. .1664)
	/gene="fabA"
	complement(1134. .1664)
	/gene="fabA"
	/EC_number="4.2.1.60"
	/note="ORF_ID:b114346"
	/codon_start=1
	/transl_table=1
	/product="3-hydroxydecanoyl-[acyl-carrier-protein]
	dehydratase
	/protein_id="BAC49611.1"
	/db_xref="GI:27352609"
	/translation="MPNPDHTPQPSYTRBELSSSEGYPGNGAOLPAPMMLMD
	RISLISLGGRRGKGIHVELDIYGHWFPCHPGDDPLMPSGLDMDMOMIGYWG
	MSGSPGKRRAIGVGVECTGETTTPVQVRIVAMRMVRGRGLVGIADGRVLADGAC
	VFTADMKRVGLTKAVE"
gene	complement(1763. .2830)
CDS	/gene="b114347"
	complement(1763. .2830)
	/gene="b114347"
	/note="ORF_ID:b114347
	hypothetical protein"
	/codon_start=1
	/transl_table=1
	/protein_id="BAC49612.1"
	/db_xref="GI:27352610"
	/translation="MKVLDHERHMAFAEVALGQIRSLRQTAIPRNYEIYVATGY
	NAPLKTITNETLARHGKLEADLEQIYETIYSHIKTTIRIDVGARVIGEIDVVKV
	GDALGWTSGSYDASGATERLSASKRQIKAVITLVASTSEMRKALIEDRLTIS
	RNEISNLQOOSLEAIRAESLTDPLTGLGNRYKFDYRMIGNAVOSALASGSLLEPLDID
	HKEISPDYVGHLTGQVRLVGLSMKQTIKQODIYARVGEEFAVLIPNTALRQALTY
	ADHIBRAVAKELIKKSTGTEILGRVTISVGSMLKPGDOTALIERADACLYAARG
	ENRVCEADPREPAVEHRSVA"
	complement(2994. .4496)
	/gene="rhlE"
	complement(2994. .4496)
	/note="ORF_ID:b114348"
	/product="ATP-dependent RNA helicase"
	/protein_id="BAC49613.1"
	/db_xref="GI:27352611"
	/translation="MSFNLGISEKVLAAVATGYTPTPIOEQAIPHYLARKVIGI
	AGTNGKAAATVLPMLTLEKGRABRPRTLLEPTREIAAOVKENDRVAGCKLN
	VALTIGVSPEDDAKLTTRGVULIATGRLLDHTERGLLITGEVLIVIDADMND
	MGFIPDIERICKLVEPTRQTLFTATWPEIRITETRELANPQKVESLPATVATVT
	OSOVPAGRKAHKEKELRLREAQDLKNAIIFCNRKKEVAIVHSLQKGSVALH
	GDMDQPARMAALEQFRKGELPLVASDVARGLDIPEVSHVNPVPHIPDDYVHVG
	RTGRGRSGTASISYTPLDOKSMVAIEKLIGOSIPRAGDVEVHAAGDATTDPREOR
	GRERSGRGKRGQRDRERBSHREAPRPSAERAPRAREPRPREARPSPEPR
	HGSRQANASHVPSIGRPERRRQREADTPEGDHSLPAPFLRPVNSPAGA"
	complement(4787. .5782)
	/gene="b114349"
	complement(4787. .5782)
	/note="rnc family
	ORF_ID:b114349"
	/codon_start=1
	/transl_table=1
	/product="transcriptional regulatory protein"
	/protein_id="BAC49614.1"
	/db_xref="GI:27352612"
	/translation="MYRWCTDVEVEQEPFYWRBYRANGLGEVTALELERERADPFE
	FSLRQLGAGLVELKASPYTVERSTSDIAYTPGDSICVYQDQSGWFGMGKSGDPAL
	ANGSFATSHTDLPYRTAPLDAGFRLRIKIVPSRIPVQDRIIRLTPTFENDHAP
	LLAACFADIGVAAADGADGADADATISVQTSIVQALADRLIERGIVRGSLGQ
	AALTRARLSOARRLIARLOLNDPLAPNAVADLIGVSHLHLFPAEKSFSQTYTDE
	RLKRRRLIRAPRELLINDIATSCGFESLATYYRVFNAYGAPDPFARVDEGH"
	5935. .7503
	/gene="b1r4350"
	5935. .7503
	/gene="b1r4350"
gene	
CDS	

```

/gene="blr4350"
/notes="ORF ID:blr4350
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC49615.1"
/db_xref="GI:27352613"
/translation="MDCTRGAGKSPSCNGNFSNTRRLFRILVTGLAALGFPALPAC
AEKRVALLVGNNDYKRVFKLKAIVNDARTMDTLKQGFVVAENQSRQSEETLLA
FDALIEPGDIAFFVAGHGEIAGQVYLLPTDVPATGEQBELVNDASILADRIVERL
ONKARTSILIVPACNNPFRKGTAVAGGGLAEMQDPEGVFSVAGSPROTLD
RNSNDADPNVSFTFRAKELQGENLVQVORTRLVSEMAITVSHRQVPPQPM
VDVFLSGAKDAARADBPPOKLAALPVSVPVRYKETTNAPIASFSRRHGASV
VSPADPTGISWRMANGDFRETGFIITLIDPTRKMRNPSTIELPDQAQITIEKY
VDOSGMQGFPIKRDPEALIRDKRILDMATSWLSFREENGILVYTHLVSRCA
IEHVRIGIDTAVPNOVLKMPGCDMDPSAITAGMPLMKLAIPATQFVSVELTYRDSV
SEIKSFRANRNN"
7987.9162
/gene="blr4351"
/987.9162
/gene="blr4351"
/notes="ORF ID:blr4351"
/codon_start=1
/transl_table=11
/product="PAD-dependent monooxygenase"
/protein_id="BAC49616.1"
/db_xref="GI:27352614"
/translation="WMLGYLLAGIDVVLVLEKADPFEDFRGDTVHPSTLOVDELG
LIDGFLKPHORLQKMDLFGGTPVRIADLSRLNTKYPIAEPQMDPFLNFRAGKR
PSLRVMTTEADVLIRGETFAGVAKRPEGIDIEADLTACDRSHSTVERGLE
VVEIGAPMDLVIRAGRKADTESVFAVPEKPEGTTDPRGQVQCAVYIAGQHDAY
KARGQALLDDVVRAPPIRSGLADYKSDIDILTVAINRLARTRRPELITGDAAH
AMSVEGVGNLAVQDAVATNLAADKLQHGCGSEDELDAVRRRRPEPRKMQVQVI
VQNNIISGALQGDRLKVLIVRLITLALPMLOGAPARLIALGVREHVSKAAPS"
9644.10108
/gene="blr4352"
/9644.10108
/gene="blr4352"
/notes="ORF ID:blr4352
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC49617.1"
/db_xref="GI:27352615"
/translation="WFGMEGTISGLDNKGSPLNTAFGAGDDVFSWRANLATVGRAG
FAVQNNLFYIKGIVAGVNNRLSVSDTVGATSGGCTHMANGMTVAGMEYGVTRNMI
VGLVYVYAFGSGTQVLOGGTTNNTFDAPKRDLOMAVVRASVKFDPVITARY"
complement(10205..10975)
/gene="kdup"
complement(10205..10975)
/gene="kdup"
/notes="ORF ID:bl14353"
/codon_start=1
/transl_table=11
/product="2-deoxy-D-gluconate 3-dehydrogenase"
/protein_id="BAC49618.1"

Query Match 8.1%; Score 113.2; DB 1; length 298700;
Best Local Similarity 54.9%; Pred. No. 1.2e-11;
Matches 267; Conservative 0; Mismatches 213; Indels 6; Gaps 2;

```

```

QY 311 CCAGCGTTGGCTTCGG---CACCAAGCCGATATGAGTGGCGTTACGATGGCTTCA 367
DB 263772 CCTCACTGGCGCTCGGCCACGTCAGAGTCCGCGCGCTTCAATCTCGCATCA 263713
QY 368 ACCGCAACCGATTAACGGTGTCCAAATGAATCTGGCATCAGATGAAGCGCAATACCCG 427
DB 263712 ACCGCAACCGATCAGCGTGTGCGCCGACCTTCATCCCGCGCTGATGAGAGATGAGG 263653
QY 428 TGGAGACGGGTAAACCGGTGCATCCGATCAAGCAGATTAATCAAAACCGGTGTGAAA 487
DB 263652 GCGACCGCTTCATCCGCTCGTTAGCGGCAAGGCGCTGCGTC---GGTGTGTGCCAGGC 263596
QY 488 AATCAACCCGAAAGGCAAGCCGTTCAATATGCGCATACCTTCCGGCGGATCTCA 547
DB 263595 GCGCAAGCTAGTGCATCCGATCCGTCGACCTTGCGGATGACCTTTCGTTCTGACCCACA 263536
QY 548 ACATCAACTGCGTTACTGCTGGCGGCTGGCGGATCAATCTGCTACTATTCGCGCG 607
DB 263535 ATTACCAATTGGCGTTCTGTATGCGCGCGCGCGCTTATCCACAGAGATGTGCGGC 263476
QY 608 CGCAGG 613
DB 263475 TGGTGG 263470

RESULT 14
AE007223/c
LOCUS AE007223 12129 bp DNA linear BCT 03-JUN-2004
DEFINITION Sinorhizobium meliloti 1021 plasmid pSyma section 29 of 121 of the
complete plasmid sequence.
ACCESSION AE007223 AE006469
VERSION AE007223.1 GI:14523386
KEYWORDS Sinorhizobium meliloti 1021
SOURCE Sinorhizobium meliloti 1021
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
REFERENCE 1 (bases 1 to 12129)
Barnett,M.J., Fisher,R.F., Jones,T., Komp,C., Abola,A.P.,
Barloy-Hubler,F., Bowser,L., Capela,D., Galibert,F., Gouzy,J.,
Gurjal,M., Hong,A., Hutzar,L., Hyman,R.W., Kahn,D., Kahn,M.L.,
Kalan,S., Keating,D.H., Palm,C., Peck,M.C., Surzycki,R.,
Wells,D.H., Yeh,K.-C., Davis,R.W., Pederspiel,N.A. and Long,S.R.
Nucleotide sequence and predicted functions of the entire
Sinorhizobium meliloti pSyma megaplasmid
Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9883-9888 (2001)
JOURNAL
MEDLINE 21396509
PUBMED 11481432
AUTHORS 2 (bases 1 to 12129)
Barnett,M.J., Fisher,R.F., Jones,T., Komp,C., Abola,A.P.,
Barloy-Hubler,F., Bowser,L., Capela,D., Galibert,F., Gouzy,J.,
Gurjal,M., Hong,A., Hutzar,L., Hyman,R.W., Kahn,D., Kahn,M.L.,
Kalan,S., Keating,D.H., Palm,C., Peck,M.C., Surzycki,R.,
Wells,D.H., Yeh,K.-C., Davis,R.W., Pederspiel,N.A. and Long,S.R.
Direct Submission
Submitted (29-MAR-2001) Biological Sciences, Stanford University,
371 Serra Mall, Stanford, CA 94305, USA
FEATURES
source
1..12129
/organism="Sinorhizobium meliloti 1021"
/mol_type="genomic DNA"
/strain="1021"
/db_xref="taxon:266834"
/plasmid="pSyma"
complement(103..318)
/locus_tag="SMA0580"
complement(103..318)
/locus_tag="SMA0580"
/function="Miscellaneous; Unknown"
/notes="Glimmer prediction"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

```

```
/protein_id="AAK64962.1"
/db_xref="GI:14523387"
/translation="MSSRALQSTVNGIGQGVAKASFRASDCSSRTACPDAAIKG
ALVGATKPPRSQTSALQADGFPVNG"
complement(394..1194)
/locus_tag="SMA0581"
complement(394..1194)
/locus_tag="SMA0581"
/function="Cell processes; Transport of small molecules"
/notes="glimmer prediction Excellent match (e-87) to
nitrate transport protein ATP binding component nasd from
Klebsiella oxytoca starting at amino acid 35 of glimmer
prediction. Met33 of glimmer prediction has good SD
sequence and was called SMA0581 start. Chrd, Nasd and Nrtc
are synonyms. Possible operon with SMA0581, SMA0583 and
SMA0585"
/codon_start=1
/transl_table=11
/product="Nitrate transport ATP binding protein, probable"
/protein_id="AAK64963.1"
/db_xref="GI:14523388"
/translation="MTKSYLSLLELDKSPERGRTREVLKQVSLTVDGKFIISIGHS
GGKSTLVIVGSLDTATTGVVLDGKVDVEGPDRAVVFQNHSLPLVTEENRLA
VKVFSRTNKKQERHEWTNRNELVQMAHAAGKPEVSGMKQVGAIRALAMEPKV
LILDEPFGALDALTRAHLQDQVQIHATTGNVLMTHDVAVLISDRIVTMTNGPS
ARVGEILDVPLARPRRIELASDRYLTGRESVLKPYERHREVEAAE"
complement(1208..2101)
/locus_tag="SMA0583"
complement(1208..2101)
/locus_tag="SMA0583"
/locus_tag="SMA0583"
/function="Cell processes; Transport of small molecules"
/notes="glimmer prediction Possible operon with SMA0581,
SMA0583 and Klebsiella (Nasb) and cyanobacteria (Nctb)"
/codon_start=1
/transl_table=11
/product="Nctb, Nitrate transport permease protein"
/protein_id="AAK64964.1"
/db_xref="GI:14523389"
/translation="MSVTNLKLRKPOATPOTAAOVIALGOTASRGIDGRSPVQTV
NILPLVLTFTPTLAKMOLICSSPESLPPRSRYLESBELIAHPPTIGGVDGQIFWH
VVASLQVALGYMAAAVAGLGTGQSLMRGLDPIEQVLRTPPLAMPLSLAA
PDGTGSAIFVIFITAIWPIIINTAVGINIQDYQNVAKVRLNSFEYFGIMLPAA
APVIFGSLRIGIGLSWALVAAEMLIGVGIGFFIWDANSSLSIDIYALVVGIVG
PLDLRLIALVGAVTRGTANA"
complement(2164..3456)
/locus_tag="SMA0585"
complement(2164..3456)
/locus_tag="SMA0585"
/locus_tag="SMA0585"
/function="Cell processes; Transport of small molecules"
/notes="glimmer prediction Excellent match (e-143) with
nitrate binding protein precursor proteins from several
cyanobacteria. Possible operon with SMA0581, SMA0583 and
SMA0585"
/codon_start=1
/transl_table=11
/product="NrtcA-type periplasmic nitrate transport binding
protein, probable"
/protein_id="AAK64965.1"
/db_xref="GI:14523390"
/translation="MTKTLFGILTRRSVLTKTATAMVGAARTLLPSGAFOAGPET
AAATLGFIALTDSAPLIIAKEKGLFDKYGTEVEVVQKASGCTRDNLVLGAGAGID
GAHLITPMPYLLISKVTQNOPLPMAIILARNLDAQAISGAAYADLKVGIIDASVLK
DAFAKKGAAGAAKVAWTFPGGTHDLWTFWMLAAGIDDPDKQVETIVVPPQMVAMK
VGTMDQFCVGEPMNEQLVUOKITGYTVNTAEIWAERPEKSPMRADWTEKPRAYKAL
VVAIEAAQWCDMDANKDELAKIVGRSWFNVPPKQIVDLRLGGEYDYGKIVENSPL
PKFPMREHASYPFOSHDAWFLTENRMGLASDTDIKGIILAVNBDIWRBAAXDLGI
SDIPASTSRGPETFDGKVPDPANPETYIKSLAISRIA"
4274..4732
/locus_tag="SMA0590"
4274..4732
/locus_tag="SMA0590"
/function="Miscellaneous; Unknown"

repeat_region
gene
CDS
/locus_tag="SMA0591"
4761..4955
/locus_tag="SMA0591"
/locus_tag="SMA0591"
/function="Miscellaneous; Hypothetical/Partial homology"
/notes="glimmer prediction Excellent match to 64 amino
acids at carboxyl terminus of much larger (44 kda) protein
from Sinorhizobium NGR234"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein, fragment"
/protein_id="AAK64967.1"
/db_xref="GI:14523392"
/translation="MVRSRWYRKLNRFRAGIEAGISGLTTRYGLAHCTRGLHHFET
YVSSVVAVNLALFARLP"
5249..6421
/locus_tag="SMA0592"
5249..6421
/locus_tag="SMA0592"
/function="Miscellaneous; Hypothetical/Globular homology"
/notes="glimmer prediction match (e-19) to two similar size
(45 kDa) proteins in Sinorhizobium sp. NGR234 over entire
length"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK64968.1"
/db_xref="GI:14523393"
/translation="MALRSANILPDEDTAGLTVETANGTRFAVHSDNNEGIACCFP
STOREHEWKVLHPFOHLGPGWMLREQASAHIVVERDILGLLRVGADICGAVSIR
PPDAQOLPRTTEATVSGRTVSGVQKLLITKQDENRFPVASAGSLIYAKPRSDR
IDNLVRNLSLRWTAAYLGEREYTGFLASLTAAYDETALVTRFDRPNBGRLEBD
CAQLLSKPKQGVAGKYDAAYEDIAAIIIRQSSRAPIDLLRFNPLIVFTLIGNDAAH
LKNPSLLETPTGLSLSPAYDVNTAFYDGPOTLALSIGERKIHLEAANOAIFFAFGK
BIGLPDRXIDQTFQQLKQVAKASIIIRPDAEPADGVHFHKEIVDNSCURILET"
6397..7428
/locus_tag="SMA0594"
6397..7428
/locus_tag="SMA0594"
/function="Miscellaneous; Unknown"
/notes="glimmer prediction"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK64969.1"
/db_xref="GI:14523394"
/translation="MPENTGNIAPLAMORYVEBALRRKAGELTQKHSLAAGVSHPT
MAAFERGGTTLTKALDILAVVGVNDEPTEGQOARVVRAPFERWRLLVPLPDSP
ARFNGWRFVPIYEGDKMLKELAFEEILKAVYRKTGWGPPFIIPREAIOPREVDL
LIECTHAPQGEVEVERGNDPACHCFMRABSGRMFLITRGYEDGAKETIPAGTIIDTL
PLMRWGEVLAELAEKLIARQADDTAVVHFRAMFTGIRGLVLSWANPLLDLVEGH
GARDEAVALEKFSFANDIESRLACMLPLTSLYERFVAGLSINRVEAEVQRLINSP
ISKREPRPR"
7698..7919
/locus_tag="SMA0596"
7698..7919
/locus_tag="SMA0596"
/function="Miscellaneous; Unknown"
```



```

/!note="glimmer prediction"
/codon_start=1
/!transl_table=11
/product="hypothetical protein"
/protein_id="AAK64970.1"
/db_xref="GI:14523395"
/!translation="MCVTICRILFLAMQWOPASTYTNMLSLRSLSCRIGLIFQINDF
DIEFATPDEGADIALAAVNGFLFKLRG"

```

```

Query Match      7.9%; Score 109.6; DB 1; Length 12129;
Best Local Similarity 52.3%; Pred. No. 6.1e-11;
Matches 402; Conservative 0; Mismatches 309; Indels 57; Gaps 5;

```

```

QY 454 ATCAAGGAGATATTATTAACCGGTCGTGAAAAATACAAAGCCGAGCCGCTTC 513
DB 2997 ATTCAGCGCTCGGTGTGAAAGATGCTTCCGCAAAAGAGCCGCTGCGCGCC 2938
QY 514 AATATGCGGATGACCTTCCCGCGGATCTACACATCAAACTGCGTTACTGCTGGCG 573
DB 2937 AAGTTCGCAATACCTTCCCGCGGATCTACACATCAAACTGCGTTACTGCTGGCG 2878
QY 574 GCTGGCGGATCAATCTGCGCTAATTCGCGCGCGGAGACATTTCCGCGCAATCGCG 633
DB 2877 GCGCGCGGATCGACCCGAGCAAGATGTCAGACATCATGCTGTGT----- 2834
QY 634 GCAGACGCTTGTGTGCGTGAACCCCGCGCGCGCAATGCGCTCCAGCTGGAAGCCGCG 693
DB 2833 -----CCGCGCGCGCGCAATGCTGCGCAATGGAAGTGTGCG 2797
QY 694 ACCATTTTGGTTATTCGCGCGCGAGCCGCTGGAACAGAGCGCGTTCGAAGGCGATA 753
DB 2796 ACCATGGAATGTTTTCGCGCGCGAGCCCTGGAAGAGCACTCTCAACGAAAGATC 2737
QY 754 GCGCGCGCGTATCAACCGATGAAGAACTCTGGAAGAGACCGCGGAAAAAGTCTTCGCG 813
DB 2736 GCGTACACCGCTGTCAACACCGCGAGATCTGGCGCGAGATCCGAGAAAGTCTTCGCG 2677
QY 814 GTGACCAACATATGGCGGAGAAATACCCCAACACTTATCTGCGGTGACCAAGCGCGTG 873
DB 2676 ATGCGCGCGGATGCGGTGAGAAAGAACCGCGCGCGCTGCAAGCGCTGTCTATGCGCATC 2617
QY 874 ATTCGCGCGCGGATCTGCGTGAACCGCAATTAACAAGACCGCAAGAGCGCATCGAA 933
DB 2616 GAGGAAGCGCGCGATGTGCGCGAG-----ATATGCGCAACAGAGCAACTCGCGCAAG 2563
QY 934 ATGCTGCGCAAAAACAATACGTCGCTGACGTCGAAGTGTGCGCGCTAGCATGAAC 993
DB 2562 ATGCTGCGCAAGCGAGCTGTTCAACGTCGCGCAAGAGACATGTCGATCGCGTGAA 2503
QY 994 GGCACCTTTGAAATACGAAAGACATTAACCGCGCGCT--ACCGACTTCAACACTTC 1050
DB 2502 GGTGAATACGATACCGCAACGCAAGATCTGAGAAAGAGCGCGCGCTTCAATGAATTC 2443
QY 1051 TTTCGCAAGCGCGCGAGCTATCGCTCTACAGAGTGAAGTGTGATCTGACCCAGCTG 1110
DB 2442 TGGCGGAAACAGCGCTCTACCGCTTCCAGAGCATGACCGCTGTTCTCAACGAAAT 2383
QY 1111 AGCGCGTGGGCGATGATCAATGAATCAACCGGCAACTGTGATCTGATACCGCGCAAG 1170
DB 2382 ATCCGCTGGGCAAGT--TGATGATGATACGAGATCAAGAGGCGTATCGCG----- 2332
QY 1171 AAGCTTACCGCGCGGACATCTATCTGCGCGCGCGCAAGAACTGCTC 1218
DB 2331 AAGTCAACCGTGAAGACATCTGCGCGGAGCGCGCAAGAGCTCTCGC 2284

```

```

RESULT 15
AP006575      299750 bp      DNA      linear      BCT 18-OCT-2003
LOCUS      Gloeobacter violaceus PCC 7421 DNA, complete genome, section 8/16.
ACCESSION      AP006575      BA000045
VERSION      AP006575.1      GI:35212526
KEYWORDS

```

```

SOURCE      Gloeobacter violaceus PCC 7421
ORGANISM      Gloeobacter violaceus PCC 7421
REFERENCE      Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
AUTHORS      1
              Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H.,
              Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y.,
              Kiyokawa, C., Kohara, M., Matsunoto, M., Matsuno, A., Nakazaki, N.,
              Shimpou, S., Takeuchi, C., Yamada, M. and Tabata, S.
              Complete genome structure of Gloeobacter violaceus PCC 7421, a
              cyanobacterium that lacks thylakoids
              DNA Res. 10, 137-145 (2003)
TITLE      2
              Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H.,
              Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y.,
              Kiyokawa, C., Kohara, M., Matsunoto, M., Matsuno, A., Nakazaki, N.,
              Shimpou, S., Takeuchi, C., Yamada, M. and Tabata, S.
              Complete genome structure of Gloeobacter violaceus PCC 7421, a
              cyanobacterium that lacks thylakoids (supplement)
              DNA Res. 10, 181-201 (2003)
JOURNAL      3 (bases 1 to 299750)
REFERENCE      Kaneko, T.
              Direct Submission
              Submitted (15-AUG-2003) Takakazu Kaneko, Kazusa DNA Research
              Institute, The First Laboratory for Plant Gene Research; 2-6-7
              Kazusa-kamatori, Kisarazu, Chiba 292-0818, Japan
              (E-mail: kaneko@kazusa.or.jp, URL: http://www.kazusa.or.jp/cyano/,
              Tel: 81-438-52-3935 (ex.2338), Fax: 81-438-52-3934)
              Location/Qualifiers
              1..299750
              /organism="Gloeobacter violaceus PCC 7421"
              /mol_type="genomic DNA"
              /strain="PCC 7421"
              /db_xref="taxon:251221"
              /complement(199..606)
              /gene="g111960"
              /complement(199..606)
              /gene="g111960"
              /note="ORF_ID:g111960
              hypothetical protein"
              /codon_start=1
              /transl_table=11
              /protein_id="BAC89901.1"
              /db_xref="GI:35212527"
              /translation="MAQRSRALNVEIYCPTGRLRMPGKHFVSTVEQAREL
              LGLTKKATLAAATTLVDRTMIEFFCPDGAAWLAACHRDKSGVSGVATGBOW
              QQTGTVDPRAPNPVSGFTYNSRGIRLRTIG"
              /complement(197..1249)
              /gene="g111961"
              /complement(197..1249)
              /gene="g111961"
              /note="ORF_ID:g111961
              hypothetical protein"
              /codon_start=1
              /transl_table=11
              /protein_id="BAC89902.1"
              /db_xref="GI:35212528"
              /translation="MKVTLLIPVYTHIDFVGHVSNIVYIOWMEIGRNKEEIGMP
              VHKLSGCIAPLIEIEIFYKSLIGDCAHWEIWLDEKSVSNVAVRMCGBEGLTV
              ATAHQGVFVDLQKRPORLPALMEREPYLOEAPABETRVABRSRK"
              /complement(1294..3900)
              /gene="g111962"
              /complement(1294..3900)
              /gene="g111962"
              /note="ORF_ID:g111962
              unknown protein"
              /codon_start=1
              /transl_table=11
              /protein_id="BAC89903.1"
              /db_xref="GI:35212529"
              /translation="WIHFGCKHRRRTWLPAVASVPCMAAOVLAIPIVPIAQCOTV
              AAPGVQOLIGMKSGASDLGDSLSPMDIAPAPQTHGAPPOPSADTQISOAPATA
              AEDASGDIIDEVSVATRRPRTERTTATYTTVTEDEPRFGLALVTVALVQVPEFOS
              LPALGVANSNGSYLRGFDPRFQRLRIGISLQRSNNRVDISRLSDIERVEVYTG

```

GATLRVAGAGAGVINTLISETPKGPBKLTAKTEAGSYGFSRYVKGGGDDTFYSYLV
YGLVAFNNPYSLTIPNTAQFYERDAVAPNGCSLFGYKPELGDVATLISGINDYIN
ASDYSXGKLVFRRDPANRLTLRLNQNONSXNEAFGPGRSAPGLCRGGTISVQPNSTSGR
RLPLDANGNELPCOTQAFIPNTPPTAPEGVRNPNANAGATVAPGOSYLIEPTGCT
TNFPQRFLOAQTEBATVDFNDYDTFTTSLNSYLSVRLTLTTTPQSPFPFDINAKGTCT
PCALTVPPPAQPTIDNQRYEQIALNSQISPGQINSFGINLEDSYQOQSGKAGQTE
FPRAISRLSFLIDISFSENELKANGLRLTNSQFGLTTPAAGVRSPPNNVLSIRA
NYSYFENPALSNLFATGSNILEANGLRPGESQVTVGFDLTPANRLGFTYPNSYL
DCAIGVSLANPNPATPAALPLVLOAQNDLTDROASGIEAVADMPTBOLRLQAIWMS
DYRAGNADNPQSTYPPYQFODBGIPENSYLLASGYVHOGMVALVNRXDGKRRRA
TTRASGVSRGSNEYPNMAATLDLNEVILPVLPFTLGSVFNLDTQYESLSGIPAPGT
TTRIGGRLEGG
/complement (3934. .6291)
/gene="g111963"
/complement (3934. .6291)
/gene="g111963"
/note="ORF_ID:g111963
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC89904.1"
/db_xref="GI:35212530"
/translation="MRRRVYLGSLILLVYGAKLAANVATKIATKPYIQPDADPG
IADQVIAQTBANGPDLYTVDFGKTPRYGRVIAGRVNHVYLSDDPLPQPSIAG
AMNHYALGLGEYETRYFCVQPGPLPEGFCASPYRKRSGAFSLVVGQGFPH
SETQRYLHYQARIIAYAMSRLEELKFPTPALPRDPFLNTGDNVYIGSGDNQYD
KMPVNSDYAACDDGAPLARSLLPIYVAGNHIDIGAGDVIINLADADAPNPAVGE
RSGALGGDALAPNNYVILPLNGPQNDPQYIPGDCRAGETPAQGTKYRBPAA
IEARYASTVPTLISGPRQIDHNSNFSFYNAHFPTDANPHELDARAGANSADK
SPPGFSYSPVSLDWLADLRSEQLMRFVYVHPAFSAITRDYQMAVARIVLEHD
GANLVENGHNHYQRTPLRAGRVAEADTVGEBEVDTAEDGIGRVPGLIHI
EGSGGFTNSGKELEDSAVGRSAAAGALPGAGWDTLTCAKMAPVYANGRGPKI
TARFKARVSPGVNVVDGDLTLTYOISEPLSATCSAADPPYSGDADGAPLRA
DLPLDPAETGLVPTALITGRLDDETVRPLDGLSTELHVPCTTGRSOTLPPRVA
LVNRSYFALNGTQVVLGPTGYTPRAIYRKEKCTQHERSVYTVGRLEBGAQKTYTIE
CVVAEAPDPVQALITASVYSTALPVAKPVTLOQGSARKTAVR"
/complement (6288. .7748)
/gene="g111964"
/complement (6288. .7748)
/gene="g111964"
/note="ORF_ID:g111964"
/codon_start=1
/transl_table=11
/product="dimethylalanine monooxygenase"
/protein_id="BAC89905.1"
/db_xref="GI:35212531"
/translation="MKRCVAYIIGGISGIVYAKCLRDGHQVTLYESIDQYGIWY
RTSGTGESVRFQNSKYLSASDYPMEQSDPPHTEILAYNSYVDHFRLRCIR
LNCQVEKRSRDMKVVSTPEGASESFDALACGVFREPRPNTIPEGADFPGLT
LHAKDYKEPSMANKRVVVMNGASGVDAVATDPAQVIFSRNSVLVRFYRAGR
PLDCNLSRYALIPAGVRKSLYRKPAPIEABHRCQNLBPAGLFPASISPAVEYDL
VAKAIEETRPITASGGRVLFITDSSSTADIVAYTAGFVSPPPHLPTESVNHGCT
DIKXVYRHPDLPRNGCFITGIIRYIGALLPCABQARFSYVLSHMOJLPTESMABEQ
NRMAQOQKQDVASGRSFQVROVEYTEETIARLIDAPKWRHMOJLAWQJLTPMLPTH
YRLDGNPKWKGSEWIRQVLDIKIPRNASGVGVR"
8491. .12084
/gene="g111965"
8491. .12084
/gene="g111965"
/note="ORF_ID:g111965"
/codon_start=1
/transl_table=11
/product="WD-repeat protein"
/protein_id="BAC89906.1"
/db_xref="GI:35212532"
/translation="MKTFRKRQRGVLLTPQGYISRLQEARRLRFRGNFNRYTLDM
SERTGLAPRTLSKVMRSKEGVDRQTLQCFHTFGLOJNEDEYLSIQIGRRRQPHAK
PMELEGAQNPFOFYREBELATLQWIIQERSRLVAVLGMGISTCTICLALQIHL
LPEEYVLRSVNGLPKELADLVAFHFPOBELPASFGEQVYSLAHAKXKGLV
VLDNETMWSGSLAGRVRRPGEYEGFLFGAVGSRDANSCLVLTSCDPOGQVPECK
RIPVRLYRIGLGIHGBRQIILBKGLAABRGPNNKELIRCTAGNPJLAKIYASTIRE
LFDGQANFLQOQSVAFGIVDLEBQHYRLSDIECLYMLAVNREPVLISQLEQDV
LLPVSRSRLLEAMESILRARSLLIECGGTFTLLPVILEYTTDLKLALVREIVDGKSLV

FRRALIKAQBPHVKAQIRFLKPLVDOLLAAVFSRLALEHYTHRLILGELRSELQ
OLGAVAGLNLILCHCGILELGGYESTLTVQADLRGLRCLNANPQNFQSITFEA
LSTVSVFSDPDQDLATSEINGITRLMQADQOALYCRGHTVMVSIASPDORLV
ASGADRPVRLMDYRTGQICLPQGBQBSVSVAFPHCGGILLASGSDAAVRMEVDS
GRCLITLRGSGWTHAVRFSPNGOVLASSQDDQKILMHPSSGEFLQMOQHTWYRS
IAFAPDQTLISGSDQDQTLRLMDVQRGLLKLQGHTRWVSVDPSADGRTLASDSD
QTVRLMDSDSLCERVMHGSNMTISVVSFSDGRLTSGSVDSYRIWISGHLRV
LQHGSGGWSVAFQDQKTLASGISIDHSVRLMDESTROPMSLQHTSMVRFVAFSPD
GTLASSGQDRTITLMDPSGRCILTKRGHWNSLAFSPNGALLASSVDHSRLIV
NVEGQCCGMIOGHTSWRSVAFPHDGRVLASQDQKTLRMDIETGRCLMTQGHTS
WRSVAFPHDQHTLASGSDQDQTVKMDVQTRLADSLSGHSGVSVFPAADGKRLAS
GGDKTVRLMDTSMQCTHVLNRRASGVLCVAIEADRSIILASSADETITLMDLQGN
YIGTMRLEGPTYGMNITGATGISEQALTLKALGAVERRDAAAPATPSGILDR"
/complement (12067. .12756)
/gene="g111966"
/complement (12067. .12756)
/gene="g111966"
/note="ORF_ID:g111966
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC89907.1"
/db_xref="GI:35212533"
/translation="MGDTGLLAVLALLTVYLAGLPLGSPALAPLVAVAGSFF
IDQFLQRLSVLVLVGLIOERRPQRRERVAHEGHLVAHRLGLPVGYYTLGAMESP
RRGQGGGGVILSGPPALGIEGESYCATVYAGLALARMYSEAVGAVEDQOQGLL
LAAQRTGAVAFRLTRNAERRAAIIDAERQTHNALAERMLTQGLERCLQLNSTK
DAASDSAVQV"
/complement (12761. .13045)
/gene="g111967"
/complement (12761. .13045)
/gene="g111967"
/note="ORF_ID:g111967
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAC89908.1"
/db_xref="GI:35212534"
/translation="MPKATNGTVLAQSDKCEVVEGNYPPPDALOTEPQPSNTRTT
CPWKGVASYVDVVDKYNKDAWYYPFPKDAAHIIITQYIAFMKOTRVEA"
/complement (13242. .14666)
/gene="g111968"
/complement (13242. .14666)
/gene="g111968"
/note="ORF_ID:g111968
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAC89909.1"
Query Match 7.5% Score 104.8; DB 1; Length 299750;
Best Local Similarity 52.2% Pred. No. 4.9e-10;
Matches 260; Conservative 0; Mismatches 232; Indels 6; Gaps 1;
QY 624 CCAATTCGCGCAGACGCTTGTGTTCGATGACCCCGCGCAATGCGGTCCAGCCT 683
DB 142467 CCACCCGACCGGACGATGATTCGACCAAGCTGCGCCGCGCATGATTCACCAACT 142526
QY 664 GGAAGCCGGCACCATTTTCGTTATTCGTCGCGAGCGCGTGGAAACGACGAGCCGTGTT 743
DB 142527 CAAGCCGGCAACATCGACGCGCTTTTGATCGGCAACCTCTGGAACATTCGCGCGCGC 142586
QY 744 CAAGGCAATAGGCGCGCGGTGATCATCCGATGAAGAACTCTGGAAAGACACGCGGAA 803
DB 142587 CGAGGCAATCGGCTTTACGTCGCGACGATCTGGAGATCTGGCCCCGTCACCCGAGCA 142646
QY 804 AGCTTCGCGGTGACCAAAACATGGCGGAGAAATATCCCAACACTTATCTGGCGGTGAC 863
DB 142647 GGTGCTGGGGGGGCGCGAGACTGGGCGCAGACTTACCCCAATATCCACTGGCCCTCAT 142706
QY 864 CAAGCGCTGATTCGGGCGCGGATTCGCTGAGCGCCGACCAATTAAGAAACCGCAAGA 923
DB 142707 CAAGCGCTCATCCAAAC-----CTGCTGTATCGCGCGAGGAAGAAAGCGCGAGGA 142760

QY 924 AGCCATCGAAATGCTGGCCGAAAAACAATACGTGGTGTGACGTGGAAGTCTGGCGGC 983
Db 142761 GATCCGCGGCTTCTTTGCCGACGAGTACTGGGGGCCGATCCGCCCTACGTCCAGTT 142820
QY 984 TAGCATGAAACGGCACTTTGGAATACGAAAAAGATAAACGGCGCTACCGGACTTCAA 1043
Db 142821 GGGCGAGGGCGAACCGAACCTGTGCCCCCTCGGCAAGAACCGCTCCAGTACGCCACCA 142880
QY 1044 CACCTTCTTTGCGCACGGCGCCAGCTATCGTCTACAGCAGTGCAGTCTGTATCTGAC 1103
Db 142881 CCTGTTGCCGCGTCCGGGGTGAACCGCCCGACCGACCGAGCAACTGTGATCTCTAC 142940
QY 1104 CCAGCTGAGGCGCTGGGG 1121
Db 142941 CCAGATGGCTCGCTGGGG 142958

Search completed: September 16, 2005, 06:53:40
Job time : 5986 secs

This Page Blank (uspto,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 02:44:35 ; Search time 751 Seconds
(without alignments)
10972.410 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392

Sequence: 1 atgaataacatcattatgac.....ccggcgcaagtcgtgat 1392

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16pec04:*
1: geneseq1808:*
2: geneseq1908:*
3: geneseq2000:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1392	100.0	1392	12	AD006542
2	143	10.3	551	12	AD006544
3	85.4	6.1	747	11	ABD14925
4	85.4	6.1	750	11	ABD15326
5	85.4	6.1	831	11	ABD15218
6	82.4	5.2	1371	11	ACH97397
7	71.8	5.2	1209	8	ACA45921
8	69	5.0	594	11	ABD15217
9	69	5.0	3720	11	ABD15023
10	62.2	4.5	1173	8	ACA24128
11	59.2	4.3	576	8	ACA23622
12	58	4.2	1200	8	ACA2316
13	58	4.2	1200	13	AD514594
14	58	4.2	1250	11	ABD10532
15	57	4.1	1116	8	ACA42316
16	50	3.6	1011	9	ACA4309
17	49.4	3.5	1512	13	ADT43829
18	48.4	3.5	773	8	ACA20473
19	48	3.4	744	13	ADT44805
20	47.8	3.4	2000	8	ADT71938

ALIGNMENTS

21	47.8	3.4	7556	8	ACA37618	ACA37618	Prokaryot
22	47.8	3.4	28804	2	AAT37329	AAT37329	Sphingomonas
23	47.8	3.4	28804	2	AAT92474	AAT92474	Sphingomonas
24	47.8	3.4	28804	2	AAV99812	AAV99812	Sphingomonas
25	47.8	3.4	28804	2	AAV81474	AAV81474	Chromosom
26	46.2	3.3	1044	12	ADJ35031	ADJ35031	DNA encod
27	46.2	3.3	37507	12	ADH48030	ADH48030	Clone PB3
28	45.6	3.3	1173	8	ACA25903	ACA25903	Prokaryot
29	45	3.2	8298	4	AAK72613	AAK72613	Human imm
30	44.8	3.2	1223	12	ADM99142	ADM99142	Environme
31	43.4	3.1	420	11	ABD15254	ABD15254	Pseudomon
32	43.4	3.1	531	11	ABD15367	ABD15367	Pseudomon
33	43.4	3.1	1119	11	ABD15153	ABD15153	Pseudomon
34	43.4	3.1	1188	8	ACA25753	ACA25753	Prokaryot
35	43.4	3.1	1389	11	ABD14772	ABD14772	Pseudomon
36	43.4	3.1	5577	13	ADT44774	ADT44774	Bacterial
37	43	3.1	1305	8	ACA23608	ACA23608	Prokaryot
38	43	3.1	1308	11	ABD03949	ABD03949	Pseudomon
39	43	3.1	1359	11	ABD03676	ABD03676	Pseudomon
40	43	3.1	1443	11	ABD04027	ABD04027	Pseudomon
41	43	3.1	1818	13	ADS48801	ADS48801	Bacterial
42	42.8	3.1	516	11	ABD10614	ABD10614	Pseudomon
43	42.8	3.1	651	11	ACH94942	ACH94942	Klebsiell
44	42.6	3.1	1149	11	ABD03368	ABD03368	Pseudomon
45	42.6	3.1	1587	11	ABD03299	ABD03299	Pseudomon

RESULT 1	AD006542	standard; DNA; 1392 BP.
ID	AD006542	
XX	AD006542	
AC	AD006542	
XX	AD006542	
DT	29-UTR-2004	(first entry)
XX	AD006542	
DE	Methylomonas sp 16a nitrogen transporter nrtA gene.	
XX	AD006542	
KW	ds; gene; C1 metabolizing bacteria; gene expression regulation;	
XX	AD006542	
OS	Methylomonas sp.; 16a.	
XX	AD006542	
FH	Key	Location/Qualifiers
FT	CDS	1..1392
FT		/*tag= a
FT		/product= "nitrogen transporter nrtA"
FT		/partial
FT		/note= "no stop codon"
XX	AD006542	
PN	MO2004037998-A2.	
XX	AD006542	
PD	06-MAY-2004.	
XX	AD006542	
PF	21-OCT-2003; 2003WO-US033698.	
XX	AD006542	
PR	21-OCT-2002; 2002US-0419872P.	
XX	AD006542	
PA	(DUFO) DU PONT DE NEMOURS & CO E I.	
XX	AD006542	
PI	Dicosimo DU, Picataggio SK, Seip JE, Ye RW, Wang T, Ni H;	
XX	AD006542	
DR	WPI: 2004-375911/35.	
XX	AD006542	
PT	Promoter regions responsive to presence of nitrate, acidic pH, elevated	
XX	AD006542	
PT	temperatures or highly expressed in presence of methane or methanol	
XX	AD006542	
PT	useful for driving expression (especially for producing zeaxanthin) in C1	
XX	AD006542	
PS	metabolizing bacteria.	
XX	AD006542	
PS	Claim 16; Page 69-70; 83pp; English.	

XX The present invention relates to a method of expressing a coding region
 CC of interest in a CI metabolizing bacteria. Also provided are protein,
 CC coding and promoter sequences from *Methylobacterium* sp. 16a which respond to
 CC various metabolic and growth conditions, which can be used in the method
 CC of the invention. The promoters are useful for driving expression of a
 CC number of coding regions (especially for producing zeaxanthin) in CI
 CC metabolizing bacteria under controlled conditions. The present sequence
 CC is a *Methylobacterium* sp. coding sequence.

XX Sequence 1392 BP; 347 A; 381 C; 405 G; 259 T; 0 U; 0 Other;

Query Match 100.0%; Score 1392; DB 12; Length 1392;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAACCATCTATGATCGAGCTCGAAGAAATTGTTATGACGTTATCGCTTGGCTA 60
 DB 1 ATGAAACCATCTATGATCGAGCTCGAAGAAATTGTTATGACGTTATCGCTTGGCTA 60
 QY 61 GCGGTTGGGGTTGAGATGCGCCGAGATGCGGGGAGTTGGCAAGCTGGAAAGAA 120
 DB 61 GCGGTTGGGGTTGAGATGCGCCGAGATGCGGGGAGTTGGCAAGCTGGAAAGAA 120
 QY 121 GATTTGAAATTCGCTTCAATCAAGCTCACCGACATGCGCGCTGGCGGTGGCCGCGAA 180
 DB 121 GATTTGAAATTCGCTTCAATCAAGCTCACCGACATGCGCGCTGGCGGTGGCCGCGAA 180
 QY 181 AAGGCTTCTTCGAGAGACGAGGCGCTGTTGTCGAATGGAAGCGGCAACTGGAAG 240
 DB 181 AAGGCTTCTTCGAGAGACGAGGCGCTGTTGTCGAATGGAAGCGGCAACTGGAAG 240
 QY 241 GTGGTATGATGAGTGGGTGCGTAATGCGAATGCGAGCGCTCGACATGCTGGCGCGCGG 300
 DB 241 GTGGTATGATGAGTGGGTGCGTAATGCGAATGCGAGCGCTCGACATGCTGGCGCGCGG 300
 QY 301 CCGTTAGCGGCGCAGCTTGGCTTCGAGCACCAGCGCGATATCGAGGTGCGCTTCAGCATG 360
 DB 301 CCGTTAGCGGCGCAGCTTGGCTTCGAGCACCAGCGCGATATCGAGGTGCGCTTCAGCATG 360
 QY 361 GGGTTCAAGCGGCAACCGGATTAACGCTGTCCATGAAATTCGGCATCAATGAAGCCGAAC 420
 DB 361 GGGTTCAAGCGGCAACCGGATTAACGCTGTCCATGAAATTCGGCATCAATGAAGCCGAAC 420
 QY 421 ATACCGCTGGAAGCGGCTAAACCGGTGATCCGATCAAGCGAGATTATCTAAACCGGTC 480
 DB 421 ATACCGCTGGAAGCGGCTAAACCGGTGATCCGATCAAGCGAGATTATCTAAACCGGTC 480
 QY 481 GTCGAATAATACAAAGCCGAGGCAAGCGGTTCAATATGCGATGACCTTCCGCGCGGA 540
 DB 481 GTCGAATAATACAAAGCCGAGGCAAGCGGTTCAATATGCGATGACCTTCCGCGCGGA 540
 QY 541 TCTCAACAATCAAACTGGCTTAATGCTGCGGCTGGGGGTATCAATCTGCTACTAT 600
 DB 541 TCTCAACAATCAAACTGGCTTAATGCTGCGGCTGGGGGTATCAATCTGCTACTAT 600
 QY 601 TCGCCGCGCAGAGACATTTCCGSCCAATCGGCGCAGACGCTTGTGTCGAGTACCCG 660
 DB 601 TCGCCGCGCAGAGACATTTCCGSCCAATCGGCGCAGACGCTTGTGTCGAGTACCCG 660
 QY 661 CCGCGCAAAATGCGCTCAACGCTGGAAGCGGCAACCATTTTGGTTATTTGCGTCCGCGAG 720
 DB 661 CCGCGCAAAATGCGCTCAACGCTGGAAGCGGCAACCATTTTGGTTATTTGCGTCCGCGAG 720
 QY 721 CCGTGAACCAAGAGCGCTGTTCAAGGCGATAGGCGTGGGTATCAACCGATGAAGAA 780
 DB 721 CCGTGAACCAAGAGCGCTGTTCAAGGCGATAGGCGTGGGTATCAACCGATGAAGAA 780
 QY 781 CTCTGGAAGGACACGCGCGGAAAAAGTCTTCGAGGTGACCAAAACATAGGCGGAGAAATAC 840
 DB 781 CTCTGGAAGGACACGCGCGGAAAAAGTCTTCGAGGTGACCAAAACATAGGCGGAGAAATAC 840
 QY 841 CCCAACCACTATCTGCGGTGACCAAGCGGCTGATTCGGGCGCGCATCTGGCTGACGCC 900

DB 841 CCCAACCACTATCTGCGGTGACCAAGCGGCTGATTCGGGCGCGCATCTGGCTGACGCC 900
 QY 901 GACATTAACAGAACCCGCAAGGAGCCATTCGAATGCTGGCGCAAAAACATATCGTGGT 960
 DB 901 GACATTAACAGAACCCGCAAGGAGCCATTCGAATGCTGGCGCAAAAACATATCGTGGT 960
 QY 961 GCTGACGTGAAGTCTGGCGGCTGACATGAAACGCGCATTTGGAATACGAAAGACGAT 1020
 DB 961 GCTGACGTGAAGTCTGGCGGCTGACATGAAACGCGCATTTGGAATACGAAAGACGAT 1020
 QY 1021 AAACGCGGCTAACCGGACTTCAACACTTCTTTCGCCACGCGCGGACATCTATCTGGC 1080
 DB 1021 AAACGCGGCTAACCGGACTTCAACACTTCTTTCGCCACGCGCGGACATCTATCTGGC 1080
 QY 1081 AGCAGTGAAGTCTGGTATCTGACCCAGCTGAGGCGCTGGGGATGATCAATGAATTCAA 1140
 DB 1081 AGCAGTGAAGTCTGGTATCTGACCCAGCTGAGGCGCTGGGGATGATCAATGAATTCAA 1140
 QY 1141 CCGGACAACTGGTATCTGGAATACCGCCAAAGACGCTACCGCGCGGACATCTATCTGGC 1200
 DB 1141 CCGGACAACTGGTATCTGGAATACCGCCAAAGACGCTACCGCGCGGACATCTATCTGGC 1200
 QY 1201 GCGGCGAAGAACTGGTGCAGAGAGGCAAGGCGCAAGGACTTCCCTGCGATAC 1260
 DB 1201 GCGGCGAAGAACTGGTGCAGAGAGGCAAGGCGCAAGGACTTCCCTGCGATAC 1260
 QY 1261 TCGATCAAGCGCTGCGCAAAATTTCTTCATCCAGCAAAAGTGCCTTGAATGCCAAGGCC 1320
 DB 1261 TCGATCAAGCGCTGCGCAAAATTTCTTCATCCAGCAAAAGTGCCTTGAATGCCAAGGCC 1320
 QY 1321 AACGATTAACCTGCGCAAGTTTGCATAGGTCTGAAGGCAAGCAACGCTAGCCGGCGC 1380
 DB 1321 AACGATTAACCTGCGCAAGTTTGCATAGGTCTGAAGGCAAGCAACGCTAGCCGGCGC 1380
 QY 1381 AAGTCTGTGAT 1392
 DB 1381 AAGTCTGTGAT 1392

RESULT 2
 AD006544
 ID AD006544 standard; DNA; 551 BP.
 XX
 AC AD006544;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE *Methylobacterium* sp. 16a nitrogen transporter *ntrA* gene promoter.
 XX
 KW de; gene; CI metabolizing bacteria; gene expression regulation;
 XX nitrogen transporter; *ntrA*; promoter.
 OS *Methylobacterium* sp.; 16a.
 XX
 PN WO2004037998-A2.
 PD 06-MAY-2004.
 XX
 PF 21-OCT-2003; 2003WO-US033698.
 XX
 PR 21-OCT-2002; 2002US-0419872P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Dicosimo DJ, Picataggio SK, Seip JE, Ye RW, Wang T, Ni H;
 DR MPI; 2004-375911/35.
 PT Promoter regions responsive to presence of nitrate, acidic pH, elevated
 PT temperatures or highly expressed in presence of methane or methanol
 PT useful for driving expression (especially for producing zeaxanthin) in CI
 PT metabolizing bacteria.

XX Claim 37; Page 72; 83pp; English.
 PS
 CC The present invention relates to a method of expressing a coding region
 CC of interest in a C1 metabolizing bacteria. Also provided are protein,
 CC coding and promoter sequences from *Methylobacterium* sp. 16a which respond to
 CC various metabolic and growth conditions, which can be used in the method
 CC of the invention. The promoters are useful for driving expression of a
 CC number of coding regions (especially for producing zeaxanthin) in C1
 CC metabolizing bacteria under controlled conditions. The present sequence
 CC is a *Methylobacterium* sp. promoter sequence.
 XX
 SQ Sequence 551 BP; 139 A; 119 C; 135 G; 158 T; 0 U; 0 Other;
 Query Match 10.3%; Score 143; DB 12; Length 551;
 Best Local Similarity 100.0%; Pred. No. 7e-27;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAAACCATTCATGAGCTCGAAGAAATTTATTGACGTTATCGGCTTCTA 60
 Db 409 ATGAAACCATTCATGAGCTCGAAGAAATTTATTGACGTTATCGGCTTCTA 468
 QY 61 GCCGTTGGGATTGACGATTCGCCGATGTCGGGCAAGTTGGCAAGTGAAGAA 120
 Db 469 GCCGTTGGGATTGACGATTCGCCGATGTCGGGCAAGTTGGCAAGTGAAGAA 528
 QY 121 GATTGAATTCGGCTTCATCA 143
 Db 529 GATTGAATTCGGCTTCATCA 551
 RESULT 3
 ABD14925/C
 ID ABD14925 standard; DNA; 747 BP.
 XX
 AC ABD14925;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE *Pseudomonas aeruginosa* polynucleotide #13529.
 XX
 KM Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection;
 KM antibacterial.
 OS *Pseudomonas aeruginosa*.
 XX
 PN US651795-B1.
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 DR MPI; 2003-615309/58.
 DR P-PSDB; ABO81354.
 XX
 PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 13529; 455pp; English.
 XX
 CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. *aeruginosa* drugs, as templates for recombinant
 CC production of P. *aeruginosa*-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. *aeruginosa*-caused
 CC infection, and in detection of P. *aeruginosa* sequences or other sequences
 CC of *Pseudomonas* species using biotech technology. Sequences ABO1397-
 CC ABD17967 represent P. *aeruginosa* polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 747 BP; 114 A; 234 C; 281 G; 118 T; 0 U; 0 Other;
 Query Match 6.1%; Score 85.4; DB 11; Length 747;
 Best Local Similarity 51.3%; Pred. No. 7.5e-12;
 Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;
 QY 651 GGTGACCCCGCCCGCCCAATGCGCTTCACGCTGGAAGCCGCGACATTTCCGTTATTG 710
 Db 559 GGTGTTTCAACCGCGCAGATGCTGCGCATTCACGCGCGCGCATCGACGCGTTCTG 500
 QY 711 CGTCGCGAGCCGCTGGAACAGAGCCGCTGTTCAAGGCGATGAGCGTGCCTGATGATC 770
 Db 499 CGCGCGCGCGCCCTTGCGGCGCCCTGCGCTGGAACAGAGCCGCGCATTCACATCGCCAC 440
 QY 771 CGATGAAGAACTCTGGAAGGACACGCGCGAAGAAAGTCTTCGCGGTGACCAACATGCGC 830
 Db 439 CAGCCAGCGATCTGCGCGGACACCGGAAAGTCTTCGCGCATACCGCGCCTTCGT 380
 QY 831 GGAAGAAATACCCCAACACTATCTGCGCGTGAACAGAGCGCTGATTCGCGCGCGCATCTG 890
 Db 379 CGAGCGCTACCCCAACACCGCGCGCGCTGCTGATGCG--GTCCTGACGCGCGCGC 323
 QY 891 GCTGACGCGCGCAATTAACAAGACCGCAAGAAAGCCATTCGAAATCTGCGCGCAAAAC 950
 Db 322 CTTCATCGAAGCAAGAGCGCGCAAGACCGCTGCGCACCGCGCACTATACGCGCGCGCA 263
 QY 951 ATACGTGCGTCTGACGTTGAAAGTCTGCGCGCTGACGATGAACGCGCATTTGCAATACGA 1010
 Db 262 CTATGTCGACGCTCGCGCTCGCGCGCATTCACGCGCGCTTCCTGCGCGCTACAGAGACG 203
 QY 1011 AAAAGCAATTAACAGCGCGCTACCGGACTTCAACACTCTTTCGCGCAAGCGCGCACTGA 1070
 Db 202 CCTCGGCAACGCGCTGCGAGACCCACATCGCTGCTTCTACGCGCGCGCAAGTCA 143
 QY 1071 TCCGTCCT--ACAGAGTGCAGTCTGATCTGACCGAGCGCTGCGCGCATGAT 1127
 Db 142 CGGCGCTGCGCTCTGCGATGCGATGCTTCAATGACCGCATTCGCGCGCTGCGCTGCT 83
 QY 1128 CAATGAA 1134
 Db 82 GCGCGAA 76
 RESULT 4
 ABD15326
 ID ABD15326 standard; DNA; 750 BP.
 XX
 AC ABD15326;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE *Pseudomonas aeruginosa* polynucleotide #13930.
 XX
 KM Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection;
 KM antibacterial.
 OS *Pseudomonas aeruginosa*.
 XX
 PN US651795-B1.
 PD 22-APR-2003.

```

XX 18-FEB-1999; 99US-00252991.
PF 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
PR 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX P-PSDB; ABO81755.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 13930; 455bp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 750 BP; 115 A; 293 C; 237 G; 105 T; 0 U; 0 Other;
XX
XX Query Match 6.1%; Score 85.4; DB 11; Length 750;
XX Best Local Similarity 51.3%; Pred. No. 7.5e-12;
XX Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;
XX
QY 651 GGTGACCCCGCCGCGCAATGCGGTCACGCTGGAAGCGGACCACTTTGCGTTATTG 710
XX |||||
DB 54 GGTGATTCCACCGCGCCAGATGTCGGCCATCTCCAGCGCGGCGATGACGGGTTCTG 113
XX |||||
QY 711 CGTGGCGAGCCGTGGAACGAGCGCGGTTCGAAGGCGATAGCGCGGTGATCAC 770
XX |||||
DB 114 CGCGGCGCGCCCTCTGGGCGCCCTTGGCCGTGACGAGGCGCAAGGCTTACCATCGCCAC 173
XX |||||
QY 771 CGATGAAGAAGCTGGAAGGACGCGGGAAGAAAGTCTTCGGCGTACCAACAAATGGGC 830
XX |||||
DB 174 CAGCCAGCGCATTTGGCCGAGCACCCGGAAAGGTCCTCGGACATACCGCGCTTCTG 233
XX |||||
QY 831 GGAGAATATCCCAACACCTATCTGGCGGTGACCAAGCGCGTGAATTCGGGCGCGATCTG 890
XX |||||
DB 234 CGAGCGCTACCCGAAACACCGCGCGCGGTGATGATGCG---GTCCTGAGACGCCAGCGCG 290
XX |||||
QY 891 GCTGAAGCGCGGACATTAACAAGACCGCAAGAACCATTCGAATCTGGCGGCAAAAACA 950
XX |||||
DB 291 CTTTCATCGAAGAGAACCGCGAGAACCGCTGGGACACCGGCAACTATCAGCGCGCCGA 350
XX |||||
QY 951 ATACGTGCGGTGAGAGTGAAGTGTGGCGGCTAGCATGAAGCGGACATTTGGAATACGA 1010
XX |||||
DB 351 CTATGTGACGCTTCGCTCGCGCGATTCAGCCGCGCTTCTTGGCGCGCTACGAGACGG 410
XX |||||
QY 1011 AAAAGAAGATTAACGCGCGCTACCGGACTTCAACACCTTCTTTCGCGACGCGCGCACTA 1070
XX |||||
DB 411 CCTCGGCAACGCGCTGACGAGACCCACATCCGCGCTTCTTCAAGCGGAGGGAATCA 470
XX |||||
QY 1071 TCCGCTCT--ACAGCATGTCAGTCTGTATCTGACCCAGCTGAGCGCTGGGGCATGAT 1127
XX |||||
DB 471 CGGCGCTGCTCGGATGTCATGTGTTCATGATCAACCGAGTTCCGCGCTGGGGCTGCT 530
XX |||||

```

```

QY 1128 CAATGAA 1134
DB 531 GCGCGAA 537
XX
XX RESULT 5
XX ABD15218
XX ID ABD15218 standard; DNA; 831 BP.
XX
XX ABD15218;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polynucleotide #13822.
XX
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX P-PSDB; ABO81647.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 13822; 455bp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 831 BP; 124 A; 327 C; 259 G; 121 T; 0 U; 0 Other;
XX
XX Query Match 6.1%; Score 85.4; DB 11; Length 831;
XX Best Local Similarity 51.3%; Pred. No. 7.8e-12;
XX Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;
XX
QY 651 GGTGACCCCGCGCGCAATGCGGTCACGCTGGAAGCGGACCACTTTGCGTTATTG 710
XX |||||
DB 100 GGTGATTCCACCGCGCCAGATGTCGGCCATCTCCAGCGCGGCGCATGACGGGTTCTG 159
XX |||||
QY 711 CGTGGCGAGCCGTGGAACGAGCGCGGTTCGAAGGCGATAGCGCGGTGATCAC 770
XX |||||
DB 160 CGCGGCGCGCCCTCTGGGCGCCCTTGGCCGTGACGAGGCGCAAGGCTTACCATCGCCAC 219
XX |||||

```


CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1209 BP; 229 A; 395 C; 376 G; 209 T; 0 U; 0 Other;

Query Match 5.2%; Score 71.8; DB 8; Length 1209;

Best Local Similarity 53.8%; Pred. No. 3e-08;

Matches 148; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 102 TGGCAAGCTGGAAGAAGATTGAAATTCGGCTTCACTCAAGCTACCGAATGGCGCC 161
 DB 45 TGACGCGCCGGAATGAACAGCTGAGACGCGGCTTCACTCAAGCTACCGAATGGCGCC 104
 QY 162 GCTGGCGGTGGCGCCGGAAGAGCTTCTTCGAGACGAGGCGCTTGTGTCATGGA 221
 DB 105 GCTGGGTGTGCTGCTGCCACCCAGGGGTTTGGCCAGCTTACGGCTCAAGCTGA 164
 QY 222 AGCGACGCGCAACTGGAAGGTGTGATGATAGAGGTGTAATGCGAACTGACGGCTC 281
 DB 165 AGCGACGACCTCTCTGGCGCGGCTGTGATGATGATGATGATGATGATGATGATGAT 224
 QY 282 GCACATGCTGGCGCGCGCGCTTACCGGCGCAAGCTTGGCTTGGGACCAAGCGCATAT 341
 DB 225 ACATACCTGTATGTTGATCTATGCTCCGTGAGAGCTGGGCAATGCGCGCGCGAC 284
 QY 342 CGAGTGCCTTTCAGCATGGGCTTCAACGGCAAG 376
 DB 285 CGACATGGCATCTCTATGGGCTCAACCAAGACG 319

RESULT 8

ABD15217

ID ABD15217 standard; DNA; 594 BP.

XX ABD15217;

XX 29-JUL-2004 (first entry)

XX *Pseudomonas aeruginosa* polynucleotide #13821.

XX Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection;

KW antibacterial.

OS *Pseudomonas aeruginosa*.

XX US651795-B1.

XX 22-APR-2003.

XX

PF 18-FEB-1999; 99US-00252991.

XX

PR 18-FEB-1998; 98US-0074788P.

PR

PR 27-JUL-1998; 98US-0094190P.

XX

XX (GENO-) GENOME THERAPEUTICS CORP.

PA

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR

DR MPI; 2003-615309/58.

DR

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.

XX

XX Disclosure; SEQ ID NO 13821; 455bp; English.

PS

XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

XX bacterial infection, for evaluating a compound, such as a polypeptide,

XX for the ability to bind a *P. aeruginosa* nucleic acid, as components of

XX effective antibacterial targets, as targets for antibacterial drugs,

XX including anti-*P. aeruginosa* drugs, as templates for recombinant

XX production of *P. aeruginosa*-derived peptides or polypeptides, as target

XX components for diagnosis and/or treatment of *P. aeruginosa*-caused

XX infection, and in detection of *P. aeruginosa* sequences or other sequences

XX of *Pseudomonas* species using biochip technology. Sequences ABD01397-

XX CC ABD17967 represent *P. aeruginosa* polynucleotides of the invention. Note:

XX The sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPRO at

XX seqdata.uspro.gov/sequence.html

XX

SEQ Sequence 594 BP; 102 A; 208 C; 196 G; 88 T; 0 U; 0 Other;

Query Match 5.0%; Score 69; DB 11; Length 594;

Best Local Similarity 52.2%; Pred. No. 1.3e-07;

Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 84 GCCCGATGTCGGGGGAGTTGGCAAGCTGGAAGAAGATTGAAATTCGGCTTCACTCA 143
 DB 204 GGCTGTGGGCGGGGCTTCCAGACGACCGAAGAGCGCGCTGACATGGTTCATGGC 263
 QY 144 GCTACCGACATGGGCGCGCTGGCGGTCGGCGCGCAAGAAAGCTTCTTGAAGAGAGG 203
 DB 264 GCTGACGATTCGGCTGCTGATGTGCGCGCCACCCAGGGCTTCGCCACCTTACGG 323
 QY 204 CTTGTTGTCGCACTGGAAGCGCAGGCAACTGGAAGGTGATGATGATGATGATGATGATGAT 263
 DB 324 CCTCAGCTCATATCTCAGGCGCAACGATCTCTGGGCGCACCTTGGCGCAACACTGCTCG 383
 QY 264 TGGCGAAGTGAAGCGCTGCGACATGCTGAGCGCGCGCGCGCTTACGGCGCGAGCTT 323
 DB 384 CGGCGAGCTGATGTCGGCGCAGTGTCTGATGCGGCTTCTTCAAGCGCTTCAAGCTGCGCT 443
 QY 324 CGGCAACGAAGCGCGATATCGAGTGCCTTCAAGCATGGGCTTCAACGGCAAG 376
 DB 444 CGGCGGACGCGCGCAAGCAGATGCGGTCTCATGGAAGCTGTGCAAAACG 496

RESULT 9

ABD15023/C

ID ABD15023 standard; DNA; 3720 BP.

XX ABD15023;

XX 29-JUL-2004 (first entry)

XX *Pseudomonas aeruginosa* polynucleotide #13627.

XX Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection;

KW

KM antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 XX
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfeld MJ, Nolling J, Deloughery C, Bush D;
 XX
 DR WPI; 2003-615309/58.
 XX
 DR P-PSDB; AB081452.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 13627; 455bp; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABD01397-
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 SO Sequence 3720 BP; 576 A; 1147 C; 1326 G; 671 T; 0 U; 0 Other;
 Query Match 5.0%; Score 69; DB 11; Length 3720;
 Best Local Similarity 52.2%; Pred. No. 2.4e-07;
 Matches 153; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
 QY 84 GCCCGATGTCGGGCGACGTTGGCAAGCTGGAAAGAAAGATTGAAATTCGGCTTATCA 143
 DB 676 GGGCTGGGCGGGGGGTTCCACCGACCGAAAGAGCGCCCTGACATCGGTTCAATGCG 617
 QY 144 GCTCACGACATGATGCGCGCGTGGCGGCGCGGCGGAAAGAGCTTCTTGAGGACGAGG 203
 DB 616 GCTGACCGATGCGCGCTCGCTGATCGTGGCGCGGCGGCGGCGGCGGCGGCGGCGG 557
 QY 204 CTTGTTCTGCAACTGGAAGCGGAGCGCACTGGAAGGTGATGATGATGATGATGATG 263
 DB 556 CTTGACGCTCAATCTCAAGCGCGCAACCGTCTGGGCGACCTTGCGGACGATGCTG 497
 QY 264 TGGCGCAATGAGAGCGGCTCGCAATGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCTT 323
 DB 496 CGGCGAGCTGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 437
 QY 324 CGGACCAAGGCGGATGATGAGGTCGCTTACGATGATGATGATGATGATGATGATGATG 376
 DB 436 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 384

XX
 AC ACA24128;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #5785.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Borrelia cepacia.
 XX
 PN WC0200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Traxwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 XX
 DR P-PSDB; AB020258.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 11998; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SO Sequence 1173 BP; 180 A; 396 C; 404 G; 193 T; 0 U; 0 Other;
 Query Match 4.5%; Score 62.2; DB 8; Length 1173;

Best Local Similarity 58.4%; Pred. No. 9.4e-06;
Matches 128; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 131 TCGGCTTCATCAAGCTCACCAGATGGCGCGCTGGCGGCGCCGCAAAAAGGCTTCT 190
 DB 116 TCGGCTATCTGCGCGATCACCGACGCGCGCCGCTGTGTCGCCCAACACGCGCTATT 175
 QY 191 TCGAGGACGAGGCGCTGT--TCGTCAACTGGAAGCGGACCACTGGAAGGTGTGA 247
 DB 176 TCGGCTCCGAAAGGCGCTGACGCTGACGACCAAGCTGCTGCTGAGCTGGGCGCAGCTCG 235
 QY 248 TGGATAGGCTGCTGAATGGCGAACTGAGCGGCTCGACATGCTGGCGCGCCGCTTAG 307
 DB 236 TCGAGGCGTTCTGTCCGCGCAAGTCAAGTCTGTGACCTGTGCTGTGGCGATGACGCTGT 295
 QY 308 CGGCGACGCTTGGCTTGGGACCAAGCGCGATTCGAGG 346
 DB 296 GGGCGGCTACGCGACCGCGCGCGCGCAAGGTGTGG 334

RESULT 11

ACA23622
ID ACA23622 standard; DNA; 576 BP.

XX ACA23622;

XX 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #5279.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

KM drug design; gene.

XX Borrelia cepacia.

OS WO20027183-A2.

XX 03-OCT-2002.

PD 21-MAR-2002; 2002MO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW,

PI Wall D, Tremick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX P-PSDB; ABU19752.

XX WPI; 2003-029926/02.

XX Claim 14; SEQ ID NO 11492; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 576 BP; 84 A; 206 C; 194 G; 92 T; 0 U; 0 Other;

Query Match 4.3%; Score 59.2; DB 8; Length 576;

Best Local Similarity 57.9%; Pred. No. 4.5e-05;
Matches 125; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 131 TCGGCTTCATCAAGCTCACCAGATGGCGCGCTGGCGGCGCCGCAAAAAGGCTTCT 190
 DB 116 TCGGCTATCTGCGCGATCACCGACGCGCGCCGCTGTGTCGCCCAACACGCGCTATT 175
 QY 191 TCGAGGACGAGGCGCTGT--TCGTCAACTGGAAGCGGACCACTGGAAGGTGTGA 247
 DB 176 TCGGCTCCGAAAGGCGCTGACGCTGACGACCAAGCTGCTGCTGAGCTGGGCGCAGCTCG 235
 QY 248 TGGATAGGCTGCTGAATGGCGAACTGAGCGGCTCGACATGCTGGCGCGCCGCTTAG 307
 DB 236 TCGAGGCGTTCTGTCCGCGCAAGTCAAGTCTGTGACCTGTGCTGTGGCGATGACGCTGT 295
 QY 308 CGGCGACGCTTGGCTTGGGACCAAGCGCGATTCGAGG 343
 DB 296 GGGCGGCTACGCGACCGCGCGCGCGCAAGGTGTGG 331

RESULT 12

ACA42316
ID ACA42316 standard; DNA; 1200 BP.

XX ACA42316;

XX 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #23973.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

KM drug design; gene.

OS Pseudomonas aeruginosa.

XX WO20027183-A2.

XX 03-OCT-2002.

PD 21-MAR-2002; 2002MO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW,

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 DR P-PSDB; AB038446.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 30186; 1766bp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1200 BP; 190 A; 420 C; 409 G; 181 T; 0 U; 0 Other;
 SQ
 Query Match 4.2%; Score 58; DB 8; Length 1200;
 Best Local Similarity 56.2%; Pred. No. 0.00012;
 Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;
 QY 131 TCGGCTTCATCAAGCTCACCGACATGCGCGCTGCGCGCGCCGCAAAAAGCTTCT 190
 DB 137 TCGGCTTACCTCGCGATCACCGACGCGCTTGTGTGGGCGCCATGCGAAGGCTGT 196
 QY 191 TCGAGACGAGGCGCGCTGTTCTGTGCACTGGAGCGGAGCC--AACTGGAGAGTGTGA 247
 DB 197 TCGAGGCGGAGGCGATTCAGGCGGAGCGCGGCTGTGCTTGTGCTTGTGCGCTCA 256
 QY 248 TGGATAGGCTGTGAATGGGGAATGAGCGGCTGCAATGTCTGCGCGCGCGCTTAG 307
 DB 257 TCGAGGCGTTCATCTCCGGCAGAGTCAAGTCACTGCTGTGTGCGCGAGTACGCTCT 316
 QY 308 CGGCGACGCTTGCTTGGGACCAAGCGCGATATCGAGTGGCGCTTACAGATGGGCTTCA 367
 DB 317 GGGCC-----CGCTACGGGAGGAGAGTGGCGGCGCAAGGTGGGCTTGAACCACTGC 370
 QY 368 ACGGCAACGAGATTACGCTGTTCATGAATC 399
 DB 371 GCGGCTCGGGGCTACGCTGTGCTCCGGAATC 402
 RESULT 13
 ADS14594
 ID ADS14594 standard; DNA; 1200 BP.

XX
 AC ADS14594;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa quorum sensing controlled gene PA2328, SEQ ID 149.
 XX
 KW Cytostatic; Anti-HIV; Antimicrobial; Immunosuppressive;
 KW quorum sensing signalling; bacterium; quorum sensing controlled gene;
 KW biofilm-associated disease; cystic fibrosis; AIDS; middle ear infection;
 KW acne; periodontal disease; gene; ds; PA2328.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO2004083385-A2.
 XX
 PD 30-SEP-2004.
 XX
 PF 11-MAR-2004; 2004WO-US007467.
 XX
 PR 14-MAR-2003; 2003US-00389647.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 XX
 PI Greenberg EP, Schuster M, Loettrich C;
 XX
 DR WPI: 2004-709932/69.
 XX
 PT Identifying a modulator of quorum sensing signalling in bacteria, useful
 PT for treating a biofilm-associated disorder, comprises contacting the cell
 PT with a quorum sensing signal molecule in the presence and absence of a
 PT test compound.
 XX
 XX Disclosure; SEQ ID NO 149; 233bp; English.
 XX
 CC The present invention relates to a method for identifying a modulator of
 CC quorum sensing signalling in bacteria. The method comprises: providing a
 CC cell that comprises a quorum sensing controlled gene (ADS14446-ADS14798),
 CC where the cell is responsive to a quorum sensing signal molecule such
 CC that a detectable signal is generated; contacting the cell with a quorum
 CC sensing signal molecule in the presence and absence of a test compound;
 CC and detecting a change in the detectable signal. The method and modulator
 CC identified by the method are useful for treating a biofilm-associated
 CC disease or disorder, e.g., cystic fibrosis, AIDS, middle ear infections,
 CC acne, periodontal disease, catheter-associated infections, and medical
 CC device-associated infections. Note: The sequence data for this patent was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1200 BP; 190 A; 420 C; 409 G; 181 T; 0 U; 0 Other;
 SQ
 Query Match 4.2%; Score 58; DB 13; Length 1200;
 Best Local Similarity 56.2%; Pred. No. 0.00012;
 Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;
 QY 131 TCGGCTTCATCAAGCTCACCGACATGCGCGCTGCGCGCGCCGCAAAAAGCTTCT 190
 DB 137 TCGGCTTACCTCGCGATCACCGACGCGCTTGTGTGGGCGCCATGCGAAGGCTGT 196
 QY 191 TCGAGACGAGGCGCGCTGTTCTGTGCACTGGAGCGGAGCC--AACTGGAGAGTGTGA 247
 DB 197 TCGAGGCGGAGGCGATTCAGGCGGAGCGCGGCTGTGCTTGTGCTTGTGCGCTCA 256
 QY 248 TGGATAGGCTGTGAATGGGGAATGAGCGGCTGCAATGTCTGCGCGCGCGCTTAG 307
 DB 257 TCGAGGCGTTCATCTCCGGCAGAGTCAAGTCACTGCTGTGTGCGCGAGTACGCTCT 316
 QY 308 CGGCGACGCTTGCTTGGGACCAAGCGCGATATCGAGTGGCGCTTACAGATGGGCTTCA 367
 DB 317 GGGCC-----CGCTACGGGAGGAGAGTGGCGGCGCAAGGTGGGCTTGAACCACTGC 370
 QY 368 ACGGCAACGAGATTACGCTGTTCATGAATC 399
 DB 371 GCGGCTCGGGGCTACGCTGTGCTCCGGAATC 402

Dh 371 GCGGCTCGGGGCTCACGGTGTCTCCGAGATC 402

RESULT 14
ABD10532
ID ABD10532 standard; DNA; 1260 BP.
XX
XX ABD10532;
AC
XX
DT 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polynucleotide #9136.
DE
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KM antibacterial.
XX
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
PN
XX 22-APR-2003.
PD
XX 18-FEB-1999; 99US-00252991.
PF
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI: 2003-615309/58.
DR P-PSDB; ABO76961.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 9136; 455pp; English.
PS
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
SQ
XX Sequence 1260 BP; 202 A; 450 C; 423 G; 185 T; 0 U; 0 Other;
Query Match 4.2%; Score 58; DB 11; Length 1260;
Best Local Similarity 56.2%; Pred. No. 0.00012;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

Dh 131 TCGGCTTCAATCAAGTCATCCGACATAGCGCGCGTGGCGCGCCGCAAAAAGCTTCT 190
XX
XX 197 TCGGCTTCAATCCGATCCGACGACCGCTTGTGTGGCCATGCCAACGGCTGT 256
Dh 191 TCGAGGACGAGGCGCTTGTGTGCACTGGAAGCGCAGGCC--AACTGGAAGTGTGTA 247
XX
XX 257 TCGAGGCGCGAGGACATCCAGGCGCAACGCGCGGTGCTGTGCTAAGTGGCGCAGGTAA 316
Dh 248 TGGATAGGCTGTGTAATGGCGAACTGACACGCTTCGACATGCTGGCGCGCGCGCTTAG 307
XX
XX

Dh 317 TCGAGGCGCTTCAATCCGAGCCAGATCAATCCATCCGCTGCTGCGGATACCGTCT 376
Qy 308 CGGCCAGCGTTCGCTTCCGACCAAGCCGATATGAGTGCCTTCAAGATGGCTTCA 367
Dh 377 GGGCC-----CGTCACGCGACAGCAAGTGTGCGCGCCGACCAAGTGTGCGGACCA 430
Qy 368 ACGGCAAGCGATATGAGTGTGCTCAATGAATC 399
Dh 431 GCGGCTCGGGGCTCACGGTGTCTCCGAGATC 462

RESULT 15
ACA44309
ID ACA44309 standard; DNA; 1116 BP.
XX
XX ACA44309;
AC
XX 19-JUN-2003 (first entry)
DT
XX
XX Prokaryotic essential gene #25966.
DE
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
XX Pseudomonas putida.
OS
XX WO200277183-A2.
PN
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-029926/02.
DR P-PSDB; ABU40439.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 32179; 1766pp; English.
PS
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 623 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-regulated gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 1116 BP; 215 A; 361 C; 342 G; 198 T; 0 U; 0 Other;

Query Match 4.1%; Score 57; DB 8; Length 1116;

Best Local Similarity 54.1%; Pred. No. 0.00021;

Matches 139; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

```
QY 131 TCGGCTTCATCAAGCTCAGCGATGAGCGCGCTGGGCGGTGGCCGCGAAAGGCTTCT 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56 TCGGCTTACTGCCATCAAGCTCAGCGCGCTGGTGGCAATTAACAAGGCTGT 115
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 TCGAGGACGAGGGCT--GTTCTGCACTGGAAGCGCAGGCCCACTGGAAGGTGTGA 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 116 TCGAGGCCGGAAGCATCAAGGCGCGCGCTGCTGCGCAGCTGGGCGAGGTGA 175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 248 TGGATAGGGTCTGTAATGGCGAATGGAAGGCTGCAATGCTGGCGCGCGCTTAG 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 176 TCGAGGCATTCTCTCGGCGCAGGTCAATGTCATCACTGCTTTCGCCGATGACGGTAT 235
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 CGGCGACGCTTGGCTTGGGCAACCAAGGCCGATATCGAAGGTGCGCTTCAGCATGGGCTTCA 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 236 GGGCGCGCTTATGGCAGCAAGGTGCCAGCCAAAGTGTGGCTGGAACCAAGTGGCGGTT 295
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 368 ACGGCAACGCGATTACG 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 296 CGGGCTTGACTGTGCG 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: September 16, 2005, 05:13:49
Job time : 759 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 04:12:19 ; Search time 4286 Seconds

(without alignments)
12362.451 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaataacatcatcagtcgac.....ccggcgcaagtcgacgac 1392

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.4	6.7	1872	9	CL982367 OaIFSC047
2	87.6	6.3	575	8	A2935061 Bf_Ba000
3	82.4	5.9	800	8	B2575941 msh2_470.
4	61.8	4.2	1085	7	CF161028 B0692F03-
5	58.6	4.4	1085	8	B2548675 pac81-60
6	49.2	3.5	1035	8	B2553077
7	48.6	3.5	1407	8	B2517273 pac81-164
8	48.4	3.5	788	8	B2550725 pac81-2775
9	47.8	3.4	459	9	CNS0052P
10	47.2	3.4	459	6	CA718866 wkm2n.pk0
11	47.2	3.4	691	6	CD894397 G118_126C
12	47.2	3.4	824	7	CD894397 G118_126C
13	46	3.3	623	6	CA699640 wkh8_pk00
14	45.6	3.3	469	6	CA066446 SCEQAD101
15	45.6	3.3	500	6	CA066446 SCEQAD101
16	45.6	3.3	500	6	CA066446 SCEQAD101
17	45.6	3.3	616	6	CA069388 SCSFAD106
18	45.6	3.3	616	6	CA069388 SCSFAD106
19	45.6	3.3	645	6	CA212608 SCSFAD111
20	45.6	3.3	645	6	CA212608 SCSFAD111
21	45.6	3.3	708	6	CA069230 SCAGRT304
22	45.6	3.3	758	6	CA069230 SCAGRT304
23	45.6	3.3	840	6	CA174317 SCFST101
24	45.6	3.3	914	6	CA106423 SCQGR101

25	45.4	3.3	660	6	CD231149	SS1_8_F06
26	45.4	3.3	663	6	CB925634	ABAI_22_G
27	45.4	3.3	811	7	CN141741	WOUNDI_1
28	45	3.2	503	8	CC351449	CGIAE24TV
29	45	3.2	824	8	B2628435	116007_9
30	45	3.2	852	8	CC031320	PUGH08TB
31	44.8	3.2	568	2	BE775653	MY-05-C-0
32	44.8	3.2	590	2	BE776036	MY-10-B-0
33	44.6	3.2	472	4	Bj210804	Bj210804
34	44.6	3.2	580	4	Bj218083	Bj218083
35	44.6	3.2	604	6	CA253508	SCILPFL410
36	44.6	3.2	686	6	CA134878	SCILPFL101
37	44.6	3.2	695	9	CL164460	msh2_358_1
38	44.6	3.2	942	8	B2575093	msh2_430.
39	44.4	3.2	930	9	CL969992	OaIFSC019
40	44.2	3.2	547	8	BH633034	1007058E1
41	44.2	3.2	672	6	CA194351	SCRFBS102
42	44.2	3.2	788	9	CC700945	OCVBN08TH
43	44.2	3.2	1483	8	B2573056	msh2_2923
44	44.2	3.2	1605	8	B2573144	msh2_2963
45	44	3.2	424	9	CL794328	OR_CBA000

ALIGNMENTS

RESULT 1
LOCUS CL982367 1872 bp DNA linear GSS 21-SEP-2004
DEFINITION OaIFSC047480 Oryza sativa Expressed Library Oryza sativa (indica
CL982367
ACCESSION CL982367 GI:52419216
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 1872)

REFERENCE
AUTHORS Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL
COMMENT
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatics
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Classes: exon-trapped.
FEATURES
source
1..1872
Location/Qualifiers
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match
Best Local Similarity 51.9%; Score 93.4; DB 9; Length 1872;
Matches 261; Conservative 0; Mismatches 236; Indels 6; Gaps 2;

QY 633 CGCAGCGCTTGTGTGCGTACCCCGCGCGGCAATGCCGTCACCGTGAAGCCGG 692
DB 552 CCATGACGCGCAGCGTGTGTGCGTCCCGCGCGGCAATGCGTGAAGTCAATCGCATGG 611
QY 693 CACCAATTTTCGTTATGTGCGCGGCGCGGCGGCAACGAGCGCGTTCAGAGGCAT 752

Db 612 CAACATGCTGGGCTTTTGGCTGCTGAGCGCTGGAACGGAGGGCGATTACGACCGCAT 671
Qy 753 AGCGCTCCGGTATACCGCATGAAGACTCTGGAAGAGACCGCGGAAAAAGCTTTCG 812
Db 672 TGGCTTCAACCGCTGCCACGTCGACATCTGGGGAGATCATCTGAAAAAATTCCTCG 731
Qy 813 CGTACCAAAACATGAGCGGAGAAATACCCCAACCTTATCTGGCGGTGACCAAGCGCT 872
Db 732 TAGCGCTCGGACTGGGTGGAGAAAAATCCGACACCGCCGGCCCTGTGAGCGCT 791
Qy 873 GATTGGGGCGCGATCTGGCTGAGCGCCGACAAATAAACAAGACCGAAGACCATCA 932
Db 792 GATGGAAGCGCGCTGATGAGGGGTCCCGGAAAAACAACG---TGAACCGCGCA 848
Qy 933 AATGCTGGCGCAAAAACATATAGTGGTGTGACGTGGAAGTCTGGCGGTATGACATGA 992
Db 849 GATCTCTTCGCGCGGCGCTCAATGCAAGAACAGATATCTACCGGGCGAATGCT 908
Qy 993 CGGCACTTTCGAATACGAAAAAGACGATTAACGGCGCTACCGGACTTCAACACTTCTT 1052
Db 909 CGGGGAATACGACACGGCTGGGGCAGCGCTGGCAGACGCGCACCCGATCCGTTCTT 968
Qy 1053 TCGCCACG---CGCCAGCTATCCGTCTACAGCACTGACGTCTGATCTGACCCAGCT 1109
Db 969 CAACGAGGGTGGCGCTGATATCCGTAACCTCTCCGACGGCATGTGTTCTTAACCAAGTT 1028
Qy 1110 GAGCGCTGGGGCATGATCAATG 1132
Db 1029 CGCGCGCTGGGCTTCTTAAG 1051

RESULT 2

AZ935061 575 bp DNA linear GSS 24-APR-2001
LOCUS BZ_Ha0003E04r B. japonicum BAC library Bradyrhizobium japonicum
DEFINITION genomic, genomic survey sequence.

ACCESSION AZ935061
VERSION AZ935061.1 GI:13777152
KEYWORDS GSS
SOURCE Bradyrhizobium japonicum
ORGANISM Bradyrhizobium japonicum
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.

REFERENCE

AUTHORS Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.,
A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
genome
1 (bases 1 to 575)

TITLE Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.,
A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
genome
1 (bases 1 to 575)

JOURNAL

MEDLINE Genome Res. 11 (8), 1434-1440 (2001)
PUBMED 21376150
11483585

COMMENT

Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Class: BAC ends
High quality sequence stop: 535.
Location/Qualifiers
1..575

FEATURES

source
/organism="Bradyrhizobium japonicum"
/mol_type="genomic DNA"
/strain="USD110"
/db_xref="taxon:375"
/lab_host="E. coli"
/clone_lib="B. japonicum BAC library"
/note="Vector: pindig0536; Site_1: HindIII"

ORIGIN

Query Match 6.3%; Score 87.6; DB 8; Length 575;

Best Local Similarity 61.1%; Pred. No. 1.5e-12;
Matches 160; Conservative 0; Mismatches 99; Indels 3; Gaps 1;
Qy 131 TGGCTTCACTCAAGCTCACCGCATGCGCGCTGGCGTGGCGCCCGGAAAAAGCTTCT 190
Db 214 TGGCTTCACTTCCGTGCTGATGCGCGCTGATGCTGCTCCATCGACAAAGGCTTTA 273
Qy 191 TCGAGACGAGAGGCGCTTTCGTGCAACTGGAAGCGCAAGCCCAACTGGAAGTGTGATG 250
Db 274 CGCGCGGAGAGGCGCTTCCAGCTGCAACTGTGCGGAGAGTCTCTGTGTCACAGTCCGG 333
Qy 251 ATAGGCTGTGTAATGCGCAATGACGCGCTGCGACATGCTGGCGCGCGCGCTTAGCGG 310
Db 334 ACAAGCTCAATATGCGCGCTGTTGACGCGCGCCCATCTGCTGGCGCGCGGCGATCGGT 393
Qy 311 CCAAGCTTGGCTTGG---CACCAAGCCGATATGAGTGCCTTTCAGCATGGCTTCA 367
Db 394 CTTCACTCGGGCTCGGCGACGTCAGAGTGGCGATGCGCGCTTCAATCTCGGCATCA 453
Qy 368 ACGGCAAGCGGATTAAGGATGTC 389
Db 454 ACGGCAAGCGGATCAAGGATGTC 475

RESULT 3

BZ575941 800 bp DNA linear GSS 17-DEC-2002
LOCUS BZ575941/c
DEFINITION msh2_470.x2 msh Pseudomonas aeruginosa genomic clone msh2_470,
genomic survey sequence.

ACCESSION BZ575941
VERSION BZ575941.1 GI:27211002
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 800)

REFERENCE

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.B., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.,
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press

JOURNAL

COMMENT Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216994
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
Location/Qualifiers
1..800
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone_lib="msh2_470"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN

Query Match 5.9%; Score 82.4; DB 8; Length 800;
Best Local Similarity 50.7%; Pred. No. 4.2e-11;
Matches 225; Conservative 0; Mismatches 216; Indels 3; Gaps 1;
Qy 651 GGTGACCCCGCGCGCAATGCGCTCCACGCTGGAAGCCGCGACCATTTTCGTTATG 710
Db 526 GGTGTTTCAACCGCGGAGATGTGCGGCATCTCCAGGCCGCGCATCGACGGTTCTG 467
Qy 711 GGTGCGGACCGGTGGAACAGAGGCGGTGTTCAAGGGCATAGGCGGTGCGGTATAC 770
Db 466 CGCGCGGAGACCTGGGCGCGCTGGCCGTGAGACAGGCGCATTCACCATCGCAC 407

QY 771 CGATGAGACTCTGGAAGGACACGCGGAAAAAGTTTGGCGTGACCAACATGGGC 830
 DB 406 CAGCCAGCGATCTGGCGGACCAACCGGAAAGTCTTGGCACTACCGCGCTTTCGT 347
 QY 831 GGAGAAATACCCCAACACTATCTGCGGTGACCAAGCGCTGATTTGGCGCGCATCTG 890
 DB 346 CGAGCGCTACCCGAAACACCGCCCGCGCTGCTCATAGC---GGTCTCGACCGCAGCCG 290
 QY 891 GCTGACGCGGACATTAACAAGACCGCAAGAAAGCCATGAAATGCTGGCGCAAAAACA 950
 DB 289 CTTTCATGAAACAGAACCGCGAGAACCGCTGGGACCGCGCAACTGATGACGCGCGCA 230
 QY 951 ATACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1010
 DB 229 CTATGTCGACGCTCCGCTCGCGCATCGACCGCGCTTTCGCGCGCTACCAAGACGG 170
 QY 1011 AAAGACGATTAACGCGCGCTACCGGACTTCAACACCTTCTTTCGCGCGCGCGCACTA 1070
 DB 169 CCTCGGCAAGCGCTGCGAGAACCGCGCATCGCGTCTTCTTATGCGACGCGCAAGTCAA 110
 QY 1071 TCCGTCCTACAGACAGTGCAGTCTG 1094
 DB 109 CCGGCGCTGCGCTCGGAGGCGTG 86

RESULT 4
 CF161028 406 bp mRNA linear EST 25-JUL-2003
 LOCUS B0692F03-5 NIA Mouse Blastocyst cDNA library (Long) Mus musculus
 DEFINITION cDNA clone NIA:B0692F03 IMAGE:30457118 5', mRNA sequence.
 ACCESSION CF161028
 VERSION CF161028.1 GI:33270577
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 406)
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method
 Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL MEDLINE
 PUBMED 21429098
 11544199

COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: B0692 row: F column: 03
 Seq primer: M13 Reverse
 High quality sequence stop: 406
 POLY-A-No.

FEATURES
 source location/Qualifiers
 1. 406
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="NIA:B0692F03 IMAGE:30457118"
 /library_type="Blastocyst"
 /dev_stage="3.5-dpc"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Blastocyst cDNA library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

extracted from a pool of 20 Blastocysts. Double-stranded cDNAs were synthesized with an oligo (dT) primer [Invitrogen: 5'-pAGCTGTTCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from 0.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by centrifugation 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-4-S. The products were purified by phenol/chloroform and centrifugation 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN
 Query Match 4.4%; Score 61.8; DB 7; Length 406;
 Best Local Similarity 52.5%; Pred. No. 1.5e-05;
 Matches 160; Conservative 0; Mismatches 142; Indels 3; Gaps 1;

QY 630 CCGCGGAGAGCGCTTGTGCGGTGACCGCGCGCAATGCGTCCAGCTGAGAGC 689
 DB 79 CGACACCGATGTCAGATCGGCGCACTGCGCGCGGAAATGTTGCGAATCTGGCGGC 138
 QY 690 CCGCACCATTTTTCGTTATTTGCGTCGCGAGCGCGTGAACAGACGCGGTTCAGAGG 749
 DB 139 CGACACCATGACAGCGGCTTCTGCGCGCGCGCGAGTGAACAGCGCGGTTCATAGCG 198
 QY 750 CATAGCGTGGCGGTGATCAGCGATGAAGACTCTGGAAGAGACCGCGGAAAAAGTCTT 809
 DB 199 CGTCGCTTCATCCACATTCGACAAAGAGATTTCTGGAAAGGCGCATCGTGGCGCGT 258
 QY 810 CCGCGTGAACAAACATGAGCGGAGAAATACCCCAACACTTCTGCGGTGACCAAGC 869
 DB 259 CGCGGCTTCAGAAATTTGTACACGCTGCGGAAACACTTGTGGAATC 318
 QY 870 GCTGAT---TCGGGCGCGCATCTGCTGAGCGCGCAATTAACAGAACCGCAAGAGC 926
 DB 319 GATCATGATGAGAGCGGCGTTCGCGCAAGGCGGAAACCGAAACAGATTGCGGAGC 378
 QY 927 CATCG 931
 DB 379 AATCG 383

RESULT 5
 BZ548675/c 1085 bp DNA linear GSS 17-DEC-2002
 LOCUS pacel1-60_1334.s1 pacel1-60 Pseudomonas aeruginosa genomic clone
 DEFINITION pacel1-60_1334, genomic survey sequence.
 ACCESSION BZ548675
 VERSION BZ548675.1 GI:27152256
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 1085)
 Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
 Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library J. Bacteriol. (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu

```

FEATURES          Class: shotgun.
                   Location/Qualifiers
    source          1..1085
                   /organism="Pseudomonas aeruginosa"
                   /mol_type="genomic DNA"
                   /strain="1-60"
                   /db_xref="taxon:287"
                   /clone="pac61-60.1334"
                   /note="clinical isolate 1-60 Whole genomic shotgun
                   library."

ORIGIN
Query Match          4.2%; Score 58.6; DB 8; Length 1085;
Best Local Similarity 52.2%; Pred. No. 0.00014;
Matches 153; Conservative 0; Mismatches 139; Indels 1; Gaps 1;

QY 84 GCCCGATGTCGGGCGAGTTGGCAAGCTGAAAAGAAAGATTGGAATTGGGCTTCATCA 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 463 GGCCTGGGCGGCGGCTTCCGACGCAACGGAAAGAGCGCCCTGGACATCGGTTTCATGGC 404

QY 144 GCTCACCGACATGGCCCGCTGGCGGTGGCCCGCGAAAGGCTTCTTGAAGACGAGG 203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 GCTGACCGATTGCGCTCGCTGATCGTGCCGCC-ACCCATGCTTGCACCGCCCTACGG 345

QY 204 CTTGTTGTCGCACTGGAAGCGCAGGCAACTGGAAGTGATGATGAGGTGCTGAA 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 CTCACAGCTCAATCTCAGGCGCCACACCTGCTGGCCACCTTGCGGACAACTGCTCAG 285

QY 264 TGGCGAAGCTGAGCGGCTCGCACATGCTGGCGCGCGCGCTTGAAGCGCCAGCGTTG 323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 CGCGAGAGCTGATGCGCGGAGTGCTGTACGGGGCTGATGAGGGGTCCAGCTCGGCT 225

QY 324 CGGCACCAAGGCGGATATCGAGTGCCCTTCAGCATGGGCTTCAAGCGCAACG 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 CGCGGCGAGCGCGCAAGCGAGATGGCGGTGCTCATGGAGCTGTGCCAAGC 172

RESULT 6
BZ563077          1035 bp      DNA      linear      GSS 17-DEC-2002
DEFINITION      pac62-164_4075.x1 pac62-164 Pseudomonas aeruginosa genomic clone
ACCESSION      BZ563077
VERSION      BZ563077.1 GI:27186268
KEYWORDS      GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM      Pseudomonas aeruginosa
               Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
               Pseudomonadaceae; Pseudomonas.
REFERENCE      1 (bases 1 to 1035)
AUTHORS      Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
               Burns,J.L., Kaul,R. and Olsen,M.V.
               Whole-Genome-Sequence variation among multiple isolates of
               Pseudomonas aeruginosa library
JOURNAL      J. Bacteriol. (2002) In press
COMMENT      Contact: Chris K. Raymond
               Genome Center
               University of Washington
               Box 352145, Seattle, WA 98105-2145, USA
               Tel: 2062216954
               Fax: 2066857244
               Email: craymond@u.washington.edu
               Class: shotgun.
FEATURES          Location/Qualifiers
    source          1..1035
                   /organism="Pseudomonas aeruginosa"
                   /mol_type="genomic DNA"
                   /strain="2-164"
                   /db_xref="taxon:287"
                   /clone="pac62-164_4075"
                   /clone_1ib="pac62-164"
                   /note="clinical isolate 2-164 Whole genomic shotgun

```

```

ORIGIN          library."
Query Match          3.5%; Score 49.2; DB 8; Length 1035;
Best Local Similarity 55.9%; Pred. No. 0.05;
Matches 113; Conservative 0; Mismatches 88; Indels 1; Gaps 1;

QY 658 CGCGCGCGCAAAATGCCGTCACAGCTGGAAGCGCGGACCATTTGGTATTATGCGTGGC 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 CCACCGGCGCACATGTCGCCATCTCCAGCGCGCGCATGACGGTTCTGCGCGGCG 507

QY 718 -GAGCCGTGAACACGACAGCGCGTTCAGGGCATAGCGTCCGATGATCAGCATGA 776
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 508 GGAACCTTGGGCGCGCTCGGCGCGTGAACAGGCGCAGGGCTTCACATCGCACGAGCA 567

QY 777 AGAATCTGGAAGGACAGCGCGGAAAAGTCTTGGCGGTGACCAACAATGCGCGAGAA 836
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 568 GGCAGATTGCGCGGCAACCGGAAAAGGCGCTCGAACTACCGCGCTTGTTGACGC 627

QY 837 ATACCCCAACCTATCTGGCG 858
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 628 CTACCCGAACACCGCGCGCG 649

RESULT 7
BZ572739          1407 bp      DNA      linear      GSS 17-DEC-2002
DEFINITION      mh2_2775.y2 mh Pseudomonas aeruginosa genomic clone mh2_2775,
ACCESSION      BZ572739
VERSION      BZ572739.1 GI:27207800
KEYWORDS      GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM      Pseudomonas aeruginosa
               Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
               Pseudomonadaceae; Pseudomonas.
REFERENCE      1 (bases 1 to 1407)
AUTHORS      Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
               Burns,J.L., Kaul,R. and Olsen,M.V.
               Whole-Genome-Sequence variation among multiple isolates of
               Pseudomonas aeruginosa library
JOURNAL      J. Bacteriol. (2002) In press
COMMENT      Contact: Chris K. Raymond
               Genome Center
               University of Washington
               Box 352145, Seattle, WA 98105-2145, USA
               Tel: 2062216954
               Fax: 2066857244
               Email: craymond@u.washington.edu
               Class: shotgun.
FEATURES          Location/Qualifiers
    source          1..1407
                   /organism="Pseudomonas aeruginosa"
                   /mol_type="genomic DNA"
                   /strain="MSH"
                   /db_xref="taxon:287"
                   /clone="mh2_2775"
                   /clone_1ib="mh"
                   /note="Environmental isolate. Whole genomic shotgun
                   library."

ORIGIN
Query Match          3.5%; Score 48.6; DB 8; Length 1407;
Best Local Similarity 50.7%; Pred. No. 0.078;
Matches 155; Conservative 0; Mismatches 148; Indels 3; Gaps 2;

QY 84 GCCCGATGTCGGGCGAGTTGGCAAGCTGAAAAGAAAGATTGGAATTGCGCTTCATCA 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 GGCCTGGGCGGCGGCTTCCGACGCAACGGAAAGAGCGCCCTGGAATCGGTTTCATGGC 389

QY 144 GCTCACCGACATGGCGCGCTG-GGGTGGCGCGCGGAAAAGGCTTCTTGAAGACGAGG 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 GCTGACCGATTGCGCTCGTGTATGTGCGCGCACACCGAGGCTTGGCCCGCTACG 449

```

```

QY      203 GCCTGTTCTGCACTGGAAGCCGCACTGGAAGTGTGATGATAGATAGGCTGTGA 262
DB      450 GCCTCAAGCTCAATCTTAGCGCCCAACCGTCTGGGCGACCTTGCGCGAACAACCTCTCA 509
QY      263 ATGGCGAAGTGAAGCGCTCGCATGCTGCGCGCGCGCTTAGCGGCGAGCG--TTGG 320
DB      510 GCGGCGACNMGATTCGNNNNCAAGTGTCTTACGGGCTTGTTACCGGCGCTCCAGCTCG 569
QY      321 CTTGGCACCAAGCGCGATATCGAGTGTCCGTTACAGATGGGCTTCAACGCGCAACCGAT 380
DB      570 CCTCGGCGGAGCGCGGACGACGAAGGCGGCTGCTCATGATGATGCGAGAACGCGAG 629
QY      381 TACGCT 386
DB      630 GCGGAT 635

RESULT 8
BZ550725/c 788 bp DNA linear GSS 17-DEC-2002
LOCUS      pacel-60_2795.y2 pacel-60 Pseudomonas aeruginosa genomic clone
DEFINITION pacel-60_2795, genomic survey sequence.
ACCESSION  BZ550725
VERSION     BZ550725.1 GI:27154306
KEYWORDS   GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM   Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE  1 (bases 1 to 788)
            Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
            Whole-Genome-Sequence Variation among multiple isolates of
            Pseudomonas aeruginosa library
            J. Bacteriol. (2002) In press
            Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Classes: shotgun.
FEATURES   Location/Qualifiers
            source          1..788
                        /organism="Pseudomonas aeruginosa"
                        /mol_type="genomic DNA"
                        /strain="1-60"
                        /db_xref="taxon:287"
                        /clone="pacel-60_2795"
                        /clone_1lb="pacel-60"
                        /note="clinical isolate 1-60 Whole genomic shotgun
                        library."
ORIGIN
Query Match      3.5%; Score 48.4; DB 8; Length 788;
Best Local Similarity 51.3%; Pred. No. 0.078;
Matches 138; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

131 TGGGCTTATCAAGCTACCGCATGCGCCGCTGCGCGTGGCCGCGCAAAAAGGTTCT 190
DB      379 TGGCTTACTGCGGATCAAGAGCGCACCGCCATGCTGTGGCCCATGCGCAACGCGCTGT 320
QY      191 TCGAGAGCAGGAGCGCTGTCTGTCGAAGTGAAGCGCAGGCAACTGGAAGTGTGATGG 250
DB      319 TCGAGACGAGGAGCATCCAGCCGAAACGCGCGTGTCTGTGGTACTGGGCGCAGGTGA 260
QY      251 ATAGGCTCTGAATGCGCAACTGGAACGCGCTCGCATGCTGGCGCGCGCGCTTAGCGG 310
DB      259 TCAAGCGCTTCAATCTCCGCTCAGAGTCAAC---GTCAATCGCGCTGTGTCGCGAGTCCCG 203
QY      311 CAAGGTTGCTTCCGACACCAAGCGCGATATGAGGTGCGTTACAGCATGGGCTTCAACG 370

```

```

DB      202 TCTGGGCGCGCTTAGCGGACAGAGTCCGCGCCAAAGGTGTGCTTGAACCACTGCGCG 143
QY      371 GCAACCGCATTTACGTTGTCCAAATGAATC 399
DB      142 GCTCGGGGCTCAACGGTGTGCTCCGAATATC 114

RESULT 9
CNS0052P 844 bp DNA linear GSS 03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR1P16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL056652
VERSION     AL056652.1 GI:4932342
KEYWORDS   GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 844)
            Direct Submision
            Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mamoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            p1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES   Location/Qualifiers
            source          1..844
                        /organism="Drosophila melanogaster"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:7227"
                        /clone="BACR1P16"
                        /clone_1lb="RPCI-98"
                        /note="end : TET3"
ORIGIN
Query Match      3.4%; Score 47.8; DB 9; Length 844;
Best Local Similarity 17.2%; Pred. No. 0.12;
Matches 60; Conservative 143; Mismatches 145; Indels 0; Gaps 0;

471 AAAACCGTCTGCAAAAATACAAAGCGAAGCAAGCGCTTCAATATGATGACGACTT 530
DB      359 AASASVSAACSSVAAASSMAAAGASCBGCAAAAMCAACGSSSSASBSRMSSCSSSAS 418
QY      531 CCGGCGCGGATCTCAACAATCAATCAATGCTTACTGCTGCGGCGCTGCGGATCAATCC 590
DB      419 CTVSSSSCSAMSSSCCTCVSSSCGMASSSCSGMSASSSSSSGAGVSVSGGAVRGRCM 478
QY      591 TGGCTACTATTCGCGCGCGAGAGACATTTCCGCCAATTCGCGCGAGACCGCTTGTTC 650
DB      479 CCSMVCCKCMCSMCCSVSVCAVCSGVSVAAGCSVGVGVRSSRRARSSGR 538
QY      651 GGTGACCCCGCGCGCAAAATCCGTCACAGCTGGAACCGCGGACATTTTCGTTATTTG 710
DB      539 GGGSSVSSGVSSSSSVGMCAACSSASVSGBSSVSVSGVSRGRCRVGGVGG 598

```

QY	711	GGTCGGCGAAGCCGTGGAACCAAGCAGCCCGTGTTCAAAGGCATATAGCGCCGCGTGAATCATC	770
Db	599	SVSSCGSSSSSSSSGCGSVRSRSRSGANGVRGSGGRGGGGGRRGGAANAPADDA	658
QY	771	CGATGAAGAACTCTGGAAGACACCCCGGAAAAAAGTCTTCGCGCTGAC	818
Db	659	ABRAKAPANAMADAMGAAVSSSVSSSBGARNGRRAAIVNVGSGVRB	706
RESULT 10			
LOCUS	CA718866	459 bp	mRNA
DEFINITION	wm2n.pk002.c3		linear
ACCESSION	CA718866		EST 26-NOV-2002
VERSION	CA718866.1		
KEYWORDS	GI:25440659		
SOURCE	EST.		
ORGANISM	Triticum aestivum (bread wheat)		
REFERENCE	Triticum aestivum		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.		
TITLE	1 (bases 1 to 459)		
JOURNAL	Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Caraher,N. and Hanafey,M.K.		
COMMENT	Dupont wheat cDNA Sequence Unpublished (2002) Contact: Scott V. Tingey Crop Genetics E. I. Dupont de Nemours and Company 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA Tel: 302-631-2602 Fax: 302-631-2607 Email: Scott.V.Tingey@USA.dupont.com Seq primer: M13.		
FEATURES			
source	location/Qualifiers		
	1..459		
	/organism="Triticum aestivum"		
	/mol_type="mRNA"		
	/cultivar="Spring wheat"		
	/db_xref="taxon:4565"		
	/clone="wm2n.pk002.c3"		
	/tissue_type="kernel"		
	/lab_host="DH10B"		
	/clone_idb="wm2n"		
	/note="Vector: pBluescript SK+, Site 1: EcoRI, Site 2: XhoI, wheat (Triticum aestivum L.) kernel malted 175 hours at 4 C, normalized"		
ORIGIN			
Query Match	3.4%;	Score 47.2;	DB 6;
Best Local Similarity	50.9%;	Pred. No. 0.15;	Length 459;
Matches 112;	Conservative 0;	Mismatches 108;	Indels 0;
		Gaps 0;	
QY	156	GGCGCGCTGGCGGTGGCGCCGCGGAAAAAGGCTTCTTCAGAGACGAGGCGCTGTTCGTGCA	215
Db	48	GGCAGCAATGGAGGCGAGCGACGAGTCCGACAAACGCGAGCGAGGTGGCCCGACCGGCGGA	107
QY	216	ACTGGAACGCGACGCGCAACTGGAAGTGTGTATGTAGTATGGGTCGTGAATGCGCAACTGGA	275
Db	108	CTCGGGGGCGTTCGCTTCGTCGCGCGTGAACGAGGAGAGGGCGCCGAGTGGAGGCGCG	167
QY	276	CGGCTCGACATGCTGTGGCGCGCGCGCGTGAAGCGGCGACGCGTGGCTTCGACACCAAGC	335
Db	168	CGGCTCGGGGGGAGAGGGGTCCGATCTCTCGGAGACGAGGGTCCGACGAGCGCAACATCCA	227
QY	336	CGATATCGAGGTGCCGTTCAAGCATGGGCTTCAACGCGAAC	375
Db	228	GGAGGCGCAGTCTCTGCTCCGCGAGGGCGCTCTCCTCAAC	267
RESULT 11			
CD894397			

LOCUS DEFINITION	CD894397	691 bp	mRNA	EST 14-JUN-2003
ACCESSION	G118.126C22F010823	G118	Triticum aestivum cDNA clone G118126C22,	
VERSION	CD894397		mRNA sequence.	
KEYWORDS	CD894397.1	GI:32665884		
SOURCE ORGANISM	EST.			
REFERENCE	Triticum aestivum (bread wheat)			
AUTHORS	Triticum aestivum			
TITLE	Euarvota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.			
JOURNAL	1 (bases 1 to 691)			
COMMENT	Genoplane. Genoplane, a major partnership french program in plant genomics Unpublished (2003) Contact: Genoplane Genoplane			
FEATURES	93. rue Henri Rochefort 91025 EVRY CEDEX France Tel: 33 1 69 47 54 00 Fax: 33 1 69 47 54 10 This sequence has been generated in the framework of the french plant genomics programme 'Genoplane' (http://www.genoplane.com and http://genoplane-info.infobiogen.fr).			
SOURCE	Location/Qualifiers 1..691 /organism="Triticum aestivum" /mol_type="mRNA" /cultivar="rectal1" /db_xref="taxon:4565" /clone="G118126C22" /tissue_type="grain (118 degrees per day after pollination)" /clone_1bp="G118"			
ORIGIN				
Query Match	3.4%;	Score 47.2;	DB 6;	Length 691;
Best Local Similarity	50.9%;	Pred. No. 0.16;		
Matches 112;	Conservative 0;	Mismatches 108;	Indels 0;	Gaps 0;
OY	156	GGCGCCCTGCGCGGTGGCGCCGCAAAAAGGCTTCTTCGAGAGACGAGGCGCTTTCGTGCA	215	
DB	171	GGCAGCAATGCGACGCGCACAAGTCGCGACAAACGGAGCGAGTGGCCCGACCGGGA	230	
OY	216	ACTGGAAGCGCAGGCCAACTGGAAGGTGTGATGATAGGCGTCGTGAATGCCAATCGA	275	
DB	231	CTCGGGGGCGCTTCCTTCGCGCGGTGAAGCGAGGAGAGGCGCCGGAATCGGAGCGG	290	
OY	276	CGGCTGCACATAGTGGCGCGCGCCCGCTTAAGCGCGAGCGTTGGCTTCGGACCAAGC	335	
DB	291	CGCTTCGCGCGCAGCGGCGTCCAGTCTCTGAGAGACGAGGCTCGACGACGCAACATCA	350	
OY	336	CGATATCGAGGTGCGCGTTCAGCATGAGGCTTCAACGCGAAC	375	
DB	351	GGAGCCGAGTCTTCGCTCCGCGAGGCGCTCTCCCTCAAC	390	
RESULT 12				
LOCUS	CN129290	824 bp	mRNA	linear EST 01-Apr-2004
DEFINITION	RHOH1_34_C09_g1_A002 Acid- and alkaline-treated roots Sorghum bicolor cDNA clone RHOH1_34_C09_A002 5', mRNA sequence.			
ACCESSION	CN129290			
VERSION	CN129290.1	GI:45956911		
KEYWORDS	EST.			
SOURCE	Sorghum bicolor (sorghum)			
ORGANISM	Sorghum bicolor			
REFERENCE	Euarvota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Sorghum.			
AUTHORS	1 (bases 1 to 824) Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C., Sun,F., Sullivan,R., Lim,S., Baetzman,A. and Pratt,L.H.			
TITLE	An EST database from Sorghum: acid- and alkaline-treated roots			

JOURNAL
COMMENT

Unpublished (2003)
 Other ESTs: RH01.34 C09.b3.A002
 Contact: Cordonnier-Fratt M.M.
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpatr@uga.edu
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
 the Human Genome Center, University of Tokyo Institute of Medical
 Science; plant material and RNA prepared at Texas A & M University;
 sequencing done in the Laboratory for Genomics and Bioinformatics,
 University of Georgia. Sequence ends have been trimmed to exclude
 vector and regions below phred quality 16. Three-prime sequences
 are presented as their reverse complement and have been trimmed to
 exclude polyA.
 Seq primer: Sug5 (CTTGTGCTCTAAAGCTGCG).

FEATURES
source

Location/Qualifiers
 1..824
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultivar="BTx623"
 /db_xref="taxon:4558"
 /clone="RH01.34 C09.A002"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /note="Acid- and alkaline-treated roots"
 /note="Organ: Root; Vector: pME18S-FL3; Site: 1: XhoI;
 Site: 2: XhoI; The library was prepared from polyA+ RNA
 from 8-day-old roots harvested from BTx623 sorghum
 seedlings grown in hydroponic culture. HCl was added to a
 pH of 3.0 to some seedlings, KOH to a pH of 9.0 for
 others. Roots were harvested 3, 12 and 27 hr after
 addition and pooled for RNA preparation. Double-stranded
 cDNA was cloned unidirectionally into different DraIII
 sites of the pME18S-FL3 vector (5-prime DraIII site is
 CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises
 the cDNA insert."

ORIGIN

Query Match 3.4%; Score 47; DB 7; Length 824;
 Best Local Similarity 52.3%; Pred. No. 0.19;
 Matches 104; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 159 GCCGCTGGCGGTCGCCGCGGCAAAAAGGCTTCTTGAGAGCAGAGGCGCTGTTGTGCACT 218
 DB 319 GCTGCTGGCGGCTTACTTCCACGACTGCTTCCGCAAGGCTGCGAGCGCGTGGTGTGCT 378
 QY 219 GGAAGCGCAGGCGCACTGGAAGGTGTGATGATAGGCTGTGAATGGCAATGGACGG 278
 DB 379 GGAAGCGCGGCGGACGAGAGAGGCGCGCGCAACCACTCGTGGGCGGCTGACGCT 438
 QY 279 CTCGCACATGCTGCGCGCGCGCGCTTAAAGCGGCGGCGGCTTGGGACCAAGGCCGA 338
 DB 439 CGTCGACGCGCGCCCAAGGCGCGCTGAGAGGCGGCTGCGCGGACCGCTCTCTCGCCGA 498
 QY 339 TATCGAGGTGCGCTTACG 357
 DB 499 CGTCGTGGCGCTGGCCACC 517

RESULT 13
 CA699640 623 bp mRNA linear EST 26-NOV-2002
 LOCUS wlk8.pk0022.g9 wlk8 Triticum aestivum cDNA clone wlk8.pk0022.g9 5'
 DEFINITION end, mRNA sequence.
 ACCESSION CA699640
 VERSION CA699640.1 GI:25421433
 KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Poideae; Triticaceae; Triticum.
 1 (bases 1 to 623)
 Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hailey,C., Yuan,Z.,
 Mao,G., Caraher,N. and Hanafey,M.K.
 Dupont Wheat cDNA Sequence
 Unpublished (2002)
 Contact: Scott V. Tingey
 Crop Genetics
 E. I. Dupont de Nemours and Company
 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@USA.dupont.com
 Seq primer: M13.
 Location/Qualifiers
 1..623
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Stephens"
 /db_xref="taxon:4565"
 /clone="w1k8.pk0022.g9"
 /issue_type="leaf"
 /clone_lib="w1k8"
 /note="Vector: pGEM-7; Site: 1: EcoRI; Site: 2: XhoI; Wheat
 (Triticum aestivum L.) seedlings 8 hr after treatment with
 6-iodo-3-propyl-2-propyloxy-4(3H)-quinazolinone"

FEATURES
source

Location/Qualifiers
 1..623
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Stephens"
 /db_xref="taxon:4565"
 /clone="w1k8.pk0022.g9"
 /issue_type="leaf"
 /clone_lib="w1k8"
 /note="Vector: pGEM-7; Site: 1: EcoRI; Site: 2: XhoI; Wheat
 (Triticum aestivum L.) seedlings 8 hr after treatment with
 6-iodo-3-propyl-2-propyloxy-4(3H)-quinazolinone"

ORIGIN

Query Match 3.3%; Score 46; DB 6; Length 623;
 Best Local Similarity 46.3%; Pred. No. 0.33;
 Matches 116; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
 QY 158 CGCGCTGGCGGTCGCCGCGCGCAAAAAGGCTTCTTGAGAGCAGAGGCGCTGTTGTGCAAC 217
 DB 51 CGAGGTGGCGGTCGAGGCGGAGGCGGAGGAGGTGACGAGATGCGCCCTGGCGGAGAGG 110
 QY 218 TGAAGCGCAGGCGCACTGGAAGGTGTGATGATAGGCTGTGAATGGCAACTGAGAG 277
 DB 111 CGAGGCGCATGACCCCGGCGGTGGCAGCAGACAGAGGCGGTGCMGATGATGTCNATGGG 170
 QY 278 GCTTCGACATGCTGGCGCGCGCGCGCTTAAAGCGGCGGCGGCTTGGCGCACCAAGCGG 337
 DB 171 GCGAGGAGCGGAGCGGATGACCCCGCGCGCGGAGGAGGCTCCAGGCGCTTACGAGCT 230
 QY 338 ATATCGAGGTGCGGCTTCAAGATGGGCTTCAACGCGCAACCGATTAACGCTGTCATGA 397
 DB 231 GACGTGTATGACGTGTGTCTCTCTGTCGTCAAGACTTAATNGAAGCGCCGACGGTG 290
 QY 398 TCTGGCATGATGAAGCCGAACATACCGCTGGAAGCGGTAAACCGGTGATC 451
 DB 291 TACCGTCAAGAGAAAGAAACATCTGCGCGCGGTTCGCGCGGTGCGCTC 344

RESULT 14
 CA066446 469 bp mRNA linear EST 23-SEP-2003
 LOCUS SCEQAD1015H11.g AD1 Saccharum officinarum cDNA clone SCEQAD1015H11
 DEFINITION 5', mRNA sequence.
 ACCESSION CA066446
 VERSION CA066446.1 GI:34917970
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

REFERENCE 1 (bases 1 to 469)
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 COMMENT Contact: Arruda P
 Centro de Biologia Molecular e Engenhariaia Genetica

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 14:00:01 ; Search time 272 Seconds
(without alignments)
8373.887 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392
Sequence: 1 atgaaacacatcatcattagatc.....ccggcggaagtcgtgcat 1392

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	85.4	6.1	747	US-09-252-991A-13529	Sequence 13529, A
2	85.4	6.1	750	US-09-252-991A-13930	Sequence 13930, A
3	85.4	6.1	831	US-09-252-991A-13822	Sequence 13822, A
4	82.4	5.9	1371	US-09-489-039A-3192	Sequence 3192, Ap
5	69	5.0	594	US-09-252-991A-13821	Sequence 13821, A
6	69	5.0	3720	US-09-252-991A-13627	Sequence 13627, A
7	58	4.2	1260	US-09-252-991A-9136	Sequence 9136, Ap
8	50	3.6	1011	US-09-328-352-3446	Sequence 3446, Ap
9	48.2	3.5	1368	US-09-302-540-7805	Sequence 7805, Ap
10	48.2	3.5	6655	US-09-302-540-777	Sequence 777, App
11	48	3.4	750	US-09-902-540-5834	Sequence 5834, Ap
12	48	3.4	72704	US-09-902-540-5834	Sequence 5834, Ap
13	47.8	3.4	28804	US-08-592-874-1	Sequence 1273, Ap
14	47.8	3.4	28804	US-08-592-874-1	Sequence 1, Appl
15	47.8	3.4	28804	US-09-096-867-2	Sequence 2, Appl
16	44.4	3.2	7218	US-08-232-463-14	Sequence 14, Appl
17	43.8	3.1	3384	US-09-902-540-3780	Sequence 3780, Ap
18	43.8	3.1	23233	US-09-902-540-1184	Sequence 1184, Ap
19	43.4	3.1	420	US-09-252-991A-13858	Sequence 13858, A
20	43.4	3.1	531	US-09-252-991A-13971	Sequence 13971, A
21	43.4	3.1	1119	US-09-252-991A-13757	Sequence 13757, A
22	43.4	3.1	1389	US-09-252-991A-13376	Sequence 13376, A
23	43.4	3.1	3523	US-09-902-540-2081	Sequence 2081, Ap
24	43.4	3.1	5577	US-09-902-540-5383	Sequence 5383, Ap
25	43.4	3.1	34552	US-09-902-540-1262	Sequence 1262, Ap
26	43	3.1	1308	US-09-252-991A-2553	Sequence 2553, Ap
27	43	3.1	1359	US-09-252-991A-2280	Sequence 2280, Ap

28	43	3.1	1443	US-09-252-991A-2631	Sequence 2631, Ap
29	42.8	3.1	516	US-09-252-991A-9218	Sequence 9218, Ap
30	42.8	3.1	651	US-09-489-039A-737	Sequence 737, App
31	42.6	3.1	1149	US-09-252-991A-1972	Sequence 1972, Ap
32	42.6	3.1	1587	US-09-252-991A-1903	Sequence 1903, Ap
33	42.6	3.1	1836	US-09-252-991A-2056	Sequence 2056, Ap
34	42.4	3.0	450	US-09-489-039A-2830	Sequence 2830, Ap
35	42.4	3.0	465	US-09-252-991A-1459	Sequence 1459, Ap
36	42.4	3.0	939	US-09-489-039A-2942	Sequence 2942, Ap
37	42.4	3.0	1512	US-09-252-991A-1551	Sequence 1551, Ap
38	42.4	3.0	1677	US-09-252-991A-1616	Sequence 1616, Ap
39	41.8	3.0	2025	US-09-252-991A-10293	Sequence 10293, A
40	41.8	3.0	2094	US-09-252-991A-10564	Sequence 10564, A
41	41.6	3.0	1425	US-09-540-236-972	Sequence 972, App
42	41.6	3.0	10528	US-09-302-540-945	Sequence 945, App
43	41.4	3.0	1164	US-09-902-540-5610	Sequence 5610, Ap
44	41.4	3.0	1206	US-09-902-540-4100	Sequence 4100, Ap
45	41.4	3.0	11812	US-09-902-540-1041	Sequence 1041, Ap

ALIGNMENTS

RESULT 1
US-09-252-991A-13529/C

Sequence 13529, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252, 991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074, 788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094, 190

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 13529

LENGTH: 747

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13529

Query Match	6.1%;	Score 85.4;	DB 4;	Length 747;
Best Local Similarity	51.3%;	Pred. No. 1.3e-12;		
Matches 250;	Conservative	0;	Mismatches 231;	Indels 6;
			Gaps	2;
QY	651	GGTGACCCCGCCGCGCAATGCGCTCCACGCTGGAAGCCCGGACCATTTTGGTTATTG	710	
DB	559	GGTGGTTCACCGCGCGCAGATGTCGCTCCAGCGCGCGGCGATCGACGCGGTTCTG	500	
QY	711	CGTCGCGGCGCGGTGGAACGAGCGCGTTCGATTAAGGCGATTAAGCGCGGTGATCAC	770	
DB	499	CGCGGCGCGCGCTTGGGCGCGCTTGGCGGTGACGAGGCGTTCAATGCCAC	440	
QY	771	CGATGAAGAACTCTGGAAGACACGCGGAAAAAGTCTTCGCGGTGACCAAAATGAGG	830	
DB	439	CAGCCAGGAGATCTGCGCGGACCAACCGGAAAAAGTCTTCGCGGTGACCAAAATGAGG	830	
QY	831	GGAGAAATATCCCAACACCTATCTGCGCGGTGACCAAGCGCGTATTCGCGCGGATCTG	890	
DB	379	CGACGCGCTATCCGAACACCGCGCGCGTGTGATGAGC---GATCTTCGACGCCCGCGG	323	
QY	891	GGTGACGCGCGCAATTAACGAAGCGGAAAGCATCGAATGCTGGCGGAAAAACA	950	
DB	322	CTTCATCGAACGAGCGCGGAAACCGCTTGGCGACCGCGCACTGATCAAGCGCGCGCA	263	
QY	951	ATACGTGCGTGTGACGTGGAAGTGTGCGGCTGACATGAACGCGCATTTTGAATACGA	1010	
DB	262	CTATGTGACGCTCGCGCTGCGCGGCGGATTCAGCGCGGCTTCTTCGCGCGGTACAGGACG	203	

QY 1011 AAAAGACGATAAAGCGCGCTACCGGACTTCAACACTTCTTTGCGCAGCGCCGAGCTA 1070
DB 202 CCGCGGCAAGCGCTGCGAGAGACCCACATCCGCTGCGTTTACGCCGAGCGCAAGTCAA 143
QY 1071 TCCGTCCT--ACAGACATGCACTGTGTATCTGACCCAGCTGAGCGCTGGGGCATGAT 1127
DB 142 CCGGCCCTGCGCTCCGAGTGCATGTGTTTCATGACCAGTTCCGCGCTGGGGCCTGCT 83
QY 1128 CAATGAA 1134
DB 82 GCGCGAA 76

RESULT 2
US-09-252-991A-13930
; Sequence 13930, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13930
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13930

Query Match 6.1%; Score 85.4; DB 4; Length 750;
Best Local Similarity 51.3%; Pred. No. 1.3e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

QY 651 GGTGACCCCGCCGCCCAATGCGCTCCAGCTGGAAGCCGCGACCATTTTGGTTATTG 710
DB 54 GGTGTTCCACCGCGCCAGATGCTGCGACATCTCCAGCGCGCGCATGACCGGTTCTG 113
QY 711 CGTCCGCGAGCCGCTGGAACAGCAGGCGCGTTCAGAGGCATAGGCGTCCGCTGATCAG 770
DB 114 CCGCGCGCGCCCTTGGGGCGCCCTGGCGCTGAGCAGGGCCAGGCTTCACTATCGCCAC 173
QY 771 CGATGAAGAACTCTGGAAGGACAGCGCGAAAAAGTCTTCGCGGTGACCAAAACATGGGC 830
DB 174 CAGCCAGCGATCTGCGCGGACCAACCGGAAAAGTCTCGGACATACCGCGCCTTCGT 233
QY 831 GGAAGAAATACCCCAACCATCTATGCGCGTGAACCAAGCGCGCTGATTCGCGCGCATCTG 890
DB 234 CAGCGCTTACCCGAAACCGCGCGCGCTGATCAATGCG--GTCCTTGACCGCCAGCGG 290
QY 891 GCTGAGCGCGCAATATACAGAACCGCAAGAGCATGGAATGCTGGCGCAAAAAA 950
DB 291 CTTTCATGAAACAGAACCGCGGAAACCGCTGGGCAACGGAATCATAGCGCGCGCA 350
QY 951 ATACGTGCTGCTGACGTGAAGAGTCTGGCGGCTAGCATGAACGCGCATTTTGAATACGA 1010
DB 351 CTATGTCGAGCTCCGCTCGGCGCGATTCACGCGCGCTTCTTCGCGCGCTACACAGACGG 410
QY 1011 AAAAGACGATAAAGCGCGCTACCGGACTTCAACACTTCTTTGCGCAGCGCCGAGCTA 1070
DB 411 CCGCGCAAGCGCTGCGAGAGCCACATCCGCTGCGTTTCTTACGCGAGCGGAAAGTCAA 470
QY 1071 TCCGTCCT--ACAGACATGCACTGTGTATCTGACCCAGCTGAGCGCTGGGGCATGAT 1127
DB 471 CCGGCCCTGCGCTCCGAGTGCATGTGTTTCATGACCAGTTCCGCGCTGGGGCCTGCT 530
QY 1128 CAATGAA 1134
|||

DB 531 GCGCGAA 537

RESULT 3
US-09-252-991A-13822
; Sequence 13822, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13822
; LENGTH: 831
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13822

Query Match 6.1%; Score 85.4; DB 4; Length 831;
Best Local Similarity 51.3%; Pred. No. 1.3e-12;
Matches 250; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

QY 651 GGTGACCCCGCCGCCCAATGCGCTCCAGCTGGAAGCCGCGACCATTTTGGTTATTG 710
DB 100 GGTGTTCCACCGCGCCAGATGCTGCGACATCTCCAGCGCGCGCATGACCGGTTCTG 159
QY 711 CGTCCGCGAGCCGCTGGAACAGCAGGCGCGTTCAGAGGCATAGGCGTCCGCTGATCAG 770
DB 160 CCGCGCGCGCCCTTGGGGCGCCCTTGGCGTGAACCGAGGCCAGGCTTCACTATCGCCAC 219
QY 771 CGATGAAGAACTCTGGAAGGACAGCGCGAAAAAGTCTTCGCGGTGACCAAAACATGGGC 830
DB 220 CAGCCAGCGATCTGCGCGGACCAACCGGAAAAGTCTTCGACATACCGCGCCTTCGT 279
QY 831 GGAAGAAATACCCCAACCATCTATCTGCGGTGAACCAAGCGCTGATTCGCGCGCATCTG 890
DB 280 CAGCGCTTACCCGAAACCGCGCGCGCTGATCAATGCG--GTCCTTGACCGCCAGCGG 336
QY 891 GCTGAGCGCGCAATATACAGAACCGCAAGAGCATGGAATGCTGGCGCAAAAAA 950
DB 337 CTTTCATGAAACGAAACCGCGGAAACCGCTGGGCAACCGCAATGATACGCGCGCA 396
QY 951 ATACGTGCTGCTGACGTGAAGAGTCTGCGCGCTGACATGAACCGGCACTTTGAAATACGA 1010
DB 397 CTATGTCGAGCTCCGCTCGGCGCGATTCAGCGCGCTTCTTCGCGCGCTACAGAGCGG 456
QY 1011 AAAAGACGATAAAGCGCGCTACCGGACTTCAACACTTCTTTGCGCAGCGCCGAGCTA 1070
DB 457 CCGCGCAAGCGCTGCGAGAGCCACATCCGCTGCGTTTCTTACGCGAGCGGAAAGTCAA 516
QY 1071 TCCGTCCT--ACAGACATGCACTGTGTATCTGACCCAGCTGAGCGCTGGGGCATGAT 1127
DB 517 CCGGCCCTGCGCTCCGAGTGCATGTGTTTCATGACCAGTTCCGCGCTGGGGCCTGCT 576
QY 1128 CAATGAA 1134
DB 577 GCGCGAA 583
|||

RESULT 4
US-09-489-039A-3192
; Sequence 3192, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Berton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

```

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3192
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-3192

```

	Query Match	5.9%	Score 82.4	IDB 4	length 1371
	Best Local Similarity	56.2%	Pred. No. 9,7e-12		
	Matches 155	Conservative 0	Mismatches 121	Indels 0	Gaps 0
QY	652	GTGACCCCGCCGCCCAATGCGCTTCACGCTGTGAAGCCGCGACACATTTTCGGTATTTC	711		
Db	694	GTGTGTCCCGCCGCCCAATGTGTGAACATGCGCATGTGGAAATATGACGCGCTTTTTC	753		
QY	712	GTGCGCGAGCGGTGGAACAGACGCGCGTGTTCAGGGCATATGCGTGCCTGTGATCAAC	771		
Db	754	GTGCGCGAGCGGTGGAACGCGCGCGCTATTAAACGACCGATCGCTTACCGCGGCACT	813		
QY	772	GATGAAGACTCTGGAAGGACACGCGCGGAAAAAGTCTTCGCGGTACCAACATATGGCG	831		
Db	814	TCCCAAGATATCTGGCGCCCGAGCATCTGGAAAAAGTGTGGGGACCCGTGCGCATCTGGGTG	873		
QY	832	GAGAAATACCCCAACACCTATCTGGCGGTGACCAAGCGCGTGATTTCGGGCGCGATCTCG	891		
Db	874	GAAAGCAACCGAACACCGCGCGCGCTGTGTTGGCGGCGCTGATGAAGCGGACGCGTGG	933		
QY	892	CTGAGCGCGCAATATGAAGACCGGCAAGAAAGCC	927		
Db	934	ATGCGCGCTTCGCGGAGAACACCCGGGAGACGCGCC	969		

```

RESULT 5
US-09-252-991A-13821
; Sequence 13821, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 31142
; SEQ ID NO 13821
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-13821

```

Query Match	5.0%	Score 69	DB 4	Length 594
Best Local Similarity	52.2%	Pred. No. 2.4e-08		
Matches 153	Conservative 0	Mismatches 140	Indels 0	Gaps 0
QY	84	GCCCGCATGTCGGGCGCATGTGGCAAGCTTGAAAGAGAAAGATTGTAATTCGGCTTCATCAA	143	
Db	204	GCGCTGGGCGCGCGGTTCCGAAGCAGCGGAAAGAGCGCCCTGGACATTCGGTTTCATGGC	263	
QY	144	GCTCACCAGCATGGCGCGCGCTGGCGGTGGCGCGCCGCAAAAAGAGCTTCTTTCAGAGCAGAGG	203	
Db	264	GCTACGCGATTTCGGCGCTGCTGATCGTGTGGCGCGCCACCAGGAGCTTTCGCCACGCCCTACGG	323	
QY	204	CTGTGTCGTGCACTCGAAGCGCAGGCCCACTCGAAGGTGTGATGATTAAGGCTGTGCA	263	

Db CCTCAGCGCTCAATCTTCAGGCGGCCAACCGTCTCTGGGCGACCTTTGGGGAGACAGTGGCTCAG 383
 QY TGGGCAACTGAGCGGCTCGCACATGCTGGCGCGCGCGCGCTTAAGCCGGACAGCTTGGCTT 322
 Db CGGCGAGCTGGAAATGCGCGCGGACAGTGGCTTAAACGGGCTTGTCTACGGGCTCAGCGCTCGGCTT 443
 QY CGGCAACCAAGGCGGATATGCGAGTGGCGGTTTCAGCATATGGGCTTCAACGGGCAAG 376
 Db CGGCGCGGCGCGCGGCAAGCGAATGGCGGTTGCTCATGGACCTTGTGCCAANAAG 496

```

RESULT 6
US-09-252-991A-13627/c
; Sequence 13627, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13627
; LENGTH: 3720
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13627

```

	Query Match	5.0%;	Score 69;	DB 4;	Length 3720;
	Best Local Similarity	52.2%;	Pred. No. 4.7e-08;		
	Matches 153;	Conservative 0;	Mismatches 140;	Indels 0;	Gaps 0;
Qy	84	GCCCGAGTGTGGGGCAATTGGCAAGCTGAGAAAAGAAAGATTGGAAATTCGGCTTCATCAA	143		
Db	676	GGCTTGGGGCCGGCGGTTCCGACGACCGGAAAAGAGGCCCTGGACATCGGTTTCATGGC	617		
Qy	144	GCTCACCGACATGGCGCCGCTGGCGGTGGCCGCCGAAAAGGCTTCTTCAGAGACGAGGG	203		
Db	616	GCTACCGGATTCGGGCTCGCTGATCGGTGGCGCGCACCGAGGGGCTTGGCCGACGGCTTCAGG	557		
Qy	204	CTGTTCGTGCACTGGAAGCGCAGGCCAACTGGAGAGTGATGATAGGTGCTGAA	263		
Db	556	CTTCACGCTCAATCTCAGCGCGCAACCGTCTTGGGGCCACTTCGCGCACAAGCTGTCTAG	497		
Qy	264	TGGGGAACCTGAACGCGCTTCGCACATGCTGGGCGCCGCGCCCTTAGCGGCCAGCGTTGGCTT	323		
Db	496	CGGCGAGCTGGAATGCCCGCAGTGCCCTGTACGGGCTTGTCTACGGCGCTTCACGCTCGGCTT	437		
Qy	324	CGGACACCAAGCGCGATATCGAGGTGCCGTTACAGATGGGCTTCAACCGGCAACG	376		
Db	436	CGGCGGACGCGCGGCAAGCAAGATGGCGGTCTCATGGGCACTGTGCCAANAACG	384		

```

RESULT 7
US-09-252-991A-9136
: Sequence 9136, Application US/09252991A
: Patent No. 6551795
:
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERGININOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9136
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-9136

```

Query Match	4.2%	Score 58;	DB 4;	Length 1260;
Best Local Similarity	56.2%	Pred. No. 2.4e-05;		
Matches 153; Conservative	0;	Mismatches 110;	Indels 9;	Gaps 2

QY 131 TCGGGCTTCAACAAGCTCAACCGACATGGCGCGCTGGGGGTGGCGCGCGCAAAAAGCTTC 190

Db 197 TCGGGTACCTGGCGAATCAACGACGCCAGCGCCTTGTCTGGTGGGCCAATGCCAAGGCGCTGT 256

QY 191 TCGAGAGCAAGGGCGCTGTTTCGTGCACCTGGAGCGCAGGCC---AACTGAGAGGTGTGA 247

Db 257 TCGAGGGCCCAAGGGCAATCAAGGCCGAAACGGCGCGGTGTGTTGCCGTAGCTGGGCCAGGTGA 316

QY 248 TCGATATAGGTCGTGAATGGCGAACTGAGACGGGCTCGACATGCTGTGGGCGCGGCGCGGTTAG 307

Db 317 TCGAGGCGTTCAATCTCCGGCAGGTACAGCTCATCACTCTGCTGTGGCCGATAAACGCTCT 376

QY 308 CGGCCAGCGTTGGCTTCCGACACCAAGCCGATATCGAGGTGCCGTTTCAGCATGGGCTTCA 367

Db 377 GGGCC-----CGCTACGGCAGCAAGTAGTGCCGCGCAAGGGTGGTGGCGCTGGAAACACAGTCG 430

QY 368 ACGCGAAGCGGATTAACGTTGTCCATGGAATC 399

Db 431 GCGGCTTCGGGGCTCAACGATGGCTCCCGAAGATTC 462

```

RESULT 8
US-09-328-352-3446
; Sequence 3446; Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: GARY L. BRECON et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3446
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3446

```

Query Match	3.6%	Score 50;	DB 4;	Length 1011;
Best Local Similarity	51.8%;	Pred. No. 0.0028;		
Matches 113; Conservative	0;	Mismatches 105;	Indels 0;	Gaps 0

[illegible]

RESULT 9

```

US-09-902-540-7805
; Sequence 7805, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIORITY FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-17, 883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 7805
; LENGTH: 1368
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-7805

```

Query Match	3.5%;	Score 48.2;	DB 4;	Length 1368;
Best Local Similarity	44.3%;	Pred. No. 0.0095;		
Matches 197; Conservative	0;	Mismatches 248;	Indels 0;	Gaps 0;

QY	195	GGACGAGGGCCTGTTCCGTGCAACTGGAAAGCCGAGGCCAACTGGAAAGTGTGATGGATG	254
Db	93	GGTCAAGTGCAGATTCAATGAACCCGGGGCGCTTCATCAAGAACCCGATGGCGCTTACAT	152
QY	255	GGTCGTGAATGAGCAACTGGAACGCGCTCCGACATGCTGGCGCCGGGCGCCGTTAGCGGCCAG	314
Db	153	CCTGGAGAAAGCCGAGCGGGAGGGGAACCTCAAGGCCCGGCGGCACCATGTTGGAGAAAC	212
QY	315	CGTTGGCTTCCGACACCAAGGCCGATATGAGGTGCGCTTCAGACATGGGCTTCAACGGCAA	374
Db	213	GTCGGGCAACACCCGACATGGCGGTGGCGCTGGCCCGCGCGTCAAAAGGCTCAACAGTGCA	272
QY	375	CGCGATTACGGTGTCCATGAAATGTGGCATCAGATGAAAGCCGAACTATCCGCTGGAAAG	434
Db	273	CTTCAACCATGCCCGAGCAAGATGTCCCTTGGAGAAAGTCAACCGCCTCAAGCGCTGGGGCG	332
QY	435	CGGTAAACCGGTGCATCCGATCAAGGACGATTTATTAACCCGGTGTGAAAAATACAA	494
Db	333	GCAGGTGGTGGTAGCGCCGAGCAAGAGTGTCGGCCGAGGACCCGGCGAGCTTACACGAGAC	392
QY	495	AGCGGAAGCAAGCCGTTCAATATGGCGATGACTTTCGCCGGCGGATCTCAACAATCAA	554
Db	393	GGCCAAAGCGCTTGCAACCGGAGACGCCCGGGCGCTTCATGTCAACCAAGTACCAACAACC	452
QY	555	ACTGCGTTACTGCGTGCGGCGTGGCGGTATCAATCTGGCTACTATTCCCGCCGAGGA	614
Db	453	CGACAACATGAGGGGCACTACAAACCAACCGGTCTGTAGATCTACAAACGACGCGAGGG	512
QY	615	CATTTCGGGCAATCGCGCAGAC	639
Db	513	CAAGTTCACACTTCTGTGGGGC	537

```

RESULT 10
US-09-902-540-777/C
; Sequence 777, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
;

```

NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 777
 ; LENGTH: 6655
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-777

Query Match 3.5%; Score 48.2; DB 4; Length 6655;
 Best Local Similarity 44.3%; Pred. No. 0.017;
 Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 195 GAGAGAGGCTCTTCTGCACTGGAAGCCCACTGGAAGGTGTGATGATG 254
 DB 4813 GGTCAAGTGGAGTTATGAACCCGGCGCTCCATCAAGACCGATGCGCTTCAAT 4754
 QY 255 GGTGTGATGAGCGAATGAGCGCTTGCACATGCTGGCGCGCGCTTAAAGCGGCG 314
 DB 4753 CCTGAGAAAGGCGAGCGGAGGGAAGCTCAAGCCCGCGGCACTCGTCAAGAAC 4694
 QY 315 CGTTGGCTTCGGACCAAGCCGATATCGAGTGGCTTCAAGATGGCTTCAAGCGCA 374
 DB 4693 GTCCGGCAACACCGGCAATGGCGCTGGCGCTGGCGCGCTCAAGGCTTCAAGTGCAT 4634
 QY 375 CGGATTAACGGTGTCCCAATGAAATCTGCAATGAGCGCAACATACCGCTGGAAG 434
 DB 4633 CTTCACCATTCGCGCAAGATGTCTCTGGAAGATCAACCGCTCAAGCGCTGGCGCG 4574
 QY 435 CGGTAAACCGGTGATCCGATCAAGGCAATTAATTAACCGGTGTGCAAAATACAA 494
 DB 4573 GCAGTGTGTGTGACCGCGCAAGCGTGGCGCGGAGACCGCGAGCTACTACGAGAC 4514
 QY 495 AGCCGAGGCAACCGCTTCAATATGGGATGATCTTCCGCGCGGATCTCAACATCAA 554
 DB 4513 GGGCAAGCGCTGACCGCGCAAGCGCGGCGGCTTCTCAACCAATACCAACACC 4454
 QY 555 ACTGCGTACTGCGCGGCTGCGGATATCAATCTGCTATTAATTCGCGCGGCGAGA 614
 DB 4453 CGACCAACATGAGGCGGCACTACCAACACCGGTCTTGAGATTAACAGACGCCAGG 4394
 QY 615 CATTCCGGCCAAATCGCGCGAGAC 639
 DB 4393 CAAGTTCGACTACTTCTGTGCGGCG 4369

RESULT 11
 US-09-902-540-5834

; Sequence 5834, Application US/09902540
 ; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 5834

; LENGTH: 750

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-5834

Query Match 3.4%; Score 48; DB 4; Length 750;
 Best Local Similarity 47.9%; Pred. No. 0.0086;
 Matches 138; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 144 GCTCACCGACATGCGCGCTGCGGCTGCGCGCGCAAAAGCTTCTTCAAGAGACGAGG 203
 DB 204 GCGCGGAGAGTTGCGCGAGGTGACGTGCTGTCAACAAAGCGGCGCTTGGGCTT 263

QY 204 CCTGTTCTGCACTGGAAGCCGAGCCCACTGGAAGGTGTGATGATGAGGCTCTGAA 263
 DB 264 GAGCTGGGCGAGTTCGCGCCCGCTTGAGAGACTGGGACAGATGTGTGACCAACGTTGA 323
 QY 264 TGGCAACTGAGCGGCTGCAATGCTGGCGCGCGCGCTTAAAGCGCGCGCTTGGCTT 323
 DB 324 GGGCTCTCTGTACTGACGACGCGGCTGCTGCGCGCGCTTAAAGCGCGCGCGCTTAA 383
 QY 324 CGGCAACCAAGCGGATATGAGGTGCGGCTTCAAGCATGAGGCTTCAACGCAACGATTA 383
 DB 384 CGTGTCAATGAATCTGCAATGATGAAGCGCAACATACCGCTTGA 431
 QY 384 GGTGTCAATGAATCTGCAATGATGAAGCGCAACATACCGCTTGA 431
 DB 444 CGGCAACCAAGCGGCTTGTGACCACTTCAAGCTTCAACCTTCCGCGCGA 491

RESULT 12

US-09-902-540-1273/C

; Sequence 1273, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 1273

; LENGTH: 72704

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(72704)

; OTHER INFORMATION: unsure at all n locations

US-09-902-540-1273

Query Match 3.4%; Score 48; DB 4; Length 72704;
 Best Local Similarity 47.9%; Pred. No. 0.046;
 Matches 138; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 144 GCTCACCGACATGCGCGCTGCGGCTGCGCGCGCAAAAGCTTCTTGAAGAGAGG 203
 DB 70204 GCGCGGAGAGTTGCGCGAGGTGAGCGTGTGTCAACAGCGGCGCTGCGGCTT 70145
 QY 204 CCTGTTCTGCACTGGAAGCCGAGCCCACTGGAAGGTGTGATGATGAGGCTCTGAA 263
 DB 70144 GAGCTGGGCGAGTTCGCGCCCGCTTGAGAGACTGGGACAGATGTGTGACCAACGTTGA 70085
 QY 264 TGGCAACTGAGCGGCTGCAATGCTGGCGCGCGCTTAAAGCGCGCGCTTGGCTT 323
 DB 70084 GGGCTCTCTGTACTGACGACGCGGCTGCTGCGCGCGCTTAAAGCGCGCGCTTAA 70025
 QY 324 CGGCAACCAAGCGGATATGAGGTGCGGCTTCAAGCATGAGGCTTCAACGCAACGATTA 383
 DB 70024 CGTGTCAATGAATCTGCAATGATGAAGCGCAACATACCGCTTGA 431
 QY 384 GGTGTCAATGAATCTGCAATGATGAAGCGCAACATACCGCTTGA 431
 DB 69964 CGGCAACCAAGCGGCTTGTGACCACTTCAAGCTTCAACCTTCCGCGCGA 69917

RESULT 13

US-08-592-874-1/C

; Sequence 1, Application US/08592874
 ; Patent No. 5854034

```

/ GENERAL INFORMATION:
/ APPLICANT: POLLOCK, THOMAS J.
/ APPLICANT: YAMAZAKI, MOTOHIDE
/ APPLICANT: THORNE, LINDA
/ APPLICANT: MIKOLAJCZAK, MARCIA
/ APPLICANT: ARMENTROUT, RICHARD W.
/ TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
/ TITLE OF INVENTION: POLYSACCHARIDE PRODUCTION
/ NUMBER OF SEQUENCES: 1
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: JULES E. GOLDBERG
/ STREET: 261 MADISON AVENUE
/ CITY: NEW YORK
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10016-2391
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/592,874
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/377,440
/ FILING DATE: 24-JAN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: GOLDBERG, JULES E.
/ REGISTRATION NUMBER: 24,408
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-986-4090
/ TELEFAX: 212-818-9479
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 28804 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
/ FRAGMENT TYPE: N-terminal
/ US-08-592-874-1

Query Match      3.4%; Score 47.8; DB 2; Length 28804;
Best Local Similarity 47.0%; Pred. No. 0.037;
Matches 148; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 64 GTTTGGGTTTGACGATGCGCCCGATGTCGGGGCAGTTGGCAAGCTGGAAAAAGAAAGAT 123
DB 18736 GTGGCGGGGCTGGCGATCAACACCGCTGGCTGGCAATCGCGCGTGGAAAGTGTGGCG 18677

QY 124 TTGAATTGGCTTCATCAAGCTCACCGACATGGCGCGCTGGCGGTGCGCGCGAATAA 183
DB 18676 CCGACGACAGGGCGCATCGCCCGATCGGCGAGACCAAGATCGTGAGTCCCGGAAGC 18617

QY 184 GGGCTTTTGAAGAGCAAGGGCCCTGTTCTGTGCAACTGGAAAGCCGAGGCCAATGGAAAGT 243
DB 18616 GGTATGTCGGGGCATCTTCTGTGGGGCGAGGGCAAAAGTTTCAAGAGGCCAGGTGTG 18557

QY 244 GTGATGATAGGGTCTGTGATGAGCGAACTGACGCGCTCCACATGCTGCGCGCGCGCGCG 303
DB 18556 ATTCAGCTCGATCCCACTATGTGGAGCGAAGCCGCCCGCGCGCGCGCGCTGTGCTC 18497

QY 304 TTAGCGGCAAGCGTTGGCTTGGCACCAAGCCGATATGAGGTGCGCTTCAGCATGGGC 363
DB 18496 AGCGCCCAAGCTCGATCCCGCGCAACCAAGCGATCATGACGCGCTGACGGAAGGGCG 18437

QY 364 TTCAACGGCAAGCGC 378
DB 18436 TTCCGCTTGGCGCG 18422

```

```

RESULT 14
US-09-096-942-2/C
/ Sequence 2, Application US/09096942
/ Patent No. 6027925
/ GENERAL INFORMATION:
/ APPLICANT: Pollock, Thomas J
/ APPLICANT: Mikolajczak, Marcia
/ APPLICANT: Yamazaki, Motohide
/ APPLICANT: Thorne, Linda
/ APPLICANT: Armentrout, Richard W
/ TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
/ TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
/ FILE REFERENCE: seq 11st for appl filed from pro. appl
/ CURRENT APPLICATION NUMBER: US/09/096,942
/ CURRENT FILING DATE: 1998-06-12
/ EARLIER APPLICATION NUMBER: 60/049,428
/ EARLIER FILING DATE: 1997-06-12
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 28804
/ TYPE: DNA
/ ORGANISM: Sphingomonas sp. S88
/ US-09-096-942-2

Query Match      3.4%; Score 47.8; DB 3; Length 28804;
Best Local Similarity 47.0%; Pred. No. 0.037;
Matches 148; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 64 GTTTGGGTTTGACGATGCGCCCGATGTCGGGGCAGTTGGCAAGCTGGAAAAAGAAAGAT 123
DB 18736 GTGGCGGGGCTGGCGATCAACACCGCTGGCTGGCAATCGCGCGTGGAAAGTGTGGCG 18677

QY 124 TTGAATTGGCTTCATCAAGCTCACCGACATGGCGCGCTGGCGGTGCGCGCGAATAA 183
DB 18676 CCGACGACAGGGCGCATCGCCCGATCGGCGAGACCAAGATCGTGAGTCCCGGAAGC 18617

QY 184 GGGCTTTTGAAGAGCAAGGGCCCTGTTCTGTGCAACTGGAAAGCCGAGGCCAATGGAAAGT 243
DB 18616 GGTATGTCGGGGCATCTTCTGTGGGGCGAGGGCAAAAGTTTCAAGAGGCCAGGTGTG 18557

QY 244 GTGATGATAGGGTCTGTGATGAGCGAACTGACGCGCTCCACATGCTGCGCGCGCGCGCG 303
DB 18556 ATTCAGCTCGATCCCACTATGTGGAGCGAAGCCGCCCGCGCGCGCGCTGTGCTC 18497

QY 304 TTAGCGGCAAGCGTTGGCTTGGCACCAAGCCGATATGAGGTGCGCTTCAGCATGGGC 363
DB 18496 AGCGCCCAAGCTCGATGCGCGCGCAACCAAGCGATCATGACGCGCTGACGGAAGGGCG 18437

QY 364 TTCAACGGCAAGCGC 378
DB 18436 TTCCGCTTGGCGCG 18422

RESULT 15
US-09-096-867-2/C
/ Sequence 2, Application US/09096867
/ Patent No. 6030817
/ GENERAL INFORMATION:
/ APPLICANT: Pollock, Thomas J
/ APPLICANT: Mikolajczak, Marcia
/ APPLICANT: Yamazaki, Motohide
/ APPLICANT: Thorne, Linda
/ APPLICANT: Armentrout, Richard W
/ TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
/ TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
/ FILE REFERENCE: seq 11st for appl filed from pro. appl
/ CURRENT APPLICATION NUMBER: US/09/096,867
/ CURRENT FILING DATE: 1998-06-11
/ EARLIER APPLICATION NUMBER: 60/049,428
/ EARLIER FILING DATE: 1997-06-12
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PatentIn Ver. 2.0

```

SEQ ID NO 2
LENGTH: 28804
TYPE: DNA
ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Query Match 3.4%; Score 47.8; DB 3; Length 28804;
Best Local Similarity 47.0%; Pred. No. 0.037;
Matches 148; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 64 GTTTGGGGTTTGAAGATTGCGCCGATGTCGGGCGAATTGCAAGCTGGAAGAAAGAAAT 123
Db 18736 GTGGCGGGGCTGGCGATCAACACCGCTGGCTGGCAATGGCCGCTGGAAGTGTGGCG 18677
QY 124 TTGAATTGGGCTTCATCAAGCTCACCGACATGGCCGCTGGCGGTGGCCCGAAGAAA 183
Db 18676 CCGACCGCAGGGCCGCAATGCCCCGATCGGCGAAGACCAAGATCGTGCAATCCCCGAAAGC 18617
QY 184 GGGTTCTTGAGAGACGAGGGCTGTTCGTGCACTGGAAGCGAGGCCAACTGGAAGTG 243
Db 18616 GGTATGTCGCGGCCATTTCTGTGGGCGAGGGCGAAGGTTCAAGAGGGCCAGTGTCTG 18557
QY 244 GTGATGATAGGGTCGTGAATGGCGAACTGACGGCTCGCACATGCTGGCGCCGCGCCG 303
Db 18556 ATCAGCGCTGATCCACCAATGTGGGACCGGAAGCCGCGAGGGCCGCTGGCTGCTC 18497
QY 304 TTAGCGGCCAGCGTTGGCTTCGGCACCAAGGCCGATATCGAGGTGCCGTTCAAGCATGGGC 363
Db 18496 AGGCGCCAGCTCGATCCGCGCAACCAAGCGATCATCGACGCTGGACGGAAGGGC 18437
QY 364 TTCAAGCGCAAGCGC 378
Db 18436 TTCCGCTTCGCGCGC 18422

Search completed: September 16, 2005, 02:48:44
Job time : 274 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 22:21:35 ; Search time 928 Seconds
(without alignments)
9993.857 Million cell updates/sec

Title: US-10-689-200-1

Perfect score: 1392

Sequence: 1 atgaataaccatcatatgagtc.....ccggcgccgaagtcgtcgat 1392

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US66_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US66_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1392	100.0	1392	19	US-10-689-200-1
2	143	10.3	551	19	US-10-689-200-3
3	71.8	5.2	1209	17	US-10-282-122A-33791
4	62.2	4.5	1173	17	US-10-282-122A-11998
5	59.2	4.3	576	17	US-10-283-122A-11492
6	58	4.2	1200	17	US-10-283-122A-10186
7	58	4.2	1200	18	US-10-389-647-149

8	57	4.1	1116	17	US-10-282-122A-32179	Sequence 32179, A
9	49.4	3.5	1512	17	US-10-369-493-42267	Sequence 42267, A
10	48.4	3.5	773	17	US-10-282-122A-8343	Sequence 8343, Ap
11	48	3.4	744	17	US-10-369-493-43243	Sequence 43243, A
12	47.8	3.4	1017	19	US-10-437-963-65118	Sequence 65118, A
13	47.8	3.4	7656	17	US-10-282-122A-25468	Sequence 25468, A
14	46.4	3.3	3048	15	US-10-156-761-3146	Sequence 3146, Ap
15	46.4	3.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
16	46.2	3.3	1440	19	US-10-437-963-82648	Sequence 82648, A
17	45.6	3.3	1173	17	US-10-282-122A-13773	Sequence 13773, A
18	45.4	3.3	1396	19	US-10-767-701-13423	Sequence 13423, A
19	44.4	3.2	930	19	US-10-437-963-97113	Sequence 97113, A
20	44.4	3.2	1365	19	US-10-437-963-56913	Sequence 56913, A
21	44	3.2	786	19	US-10-437-963-13840	Sequence 13840, A
22	43.4	3.1	1188	17	US-10-282-122A-13623	Sequence 13623, A
23	43.4	3.1	5577	17	US-10-369-493-43212	Sequence 43212, A
24	43.2	3.1	1155	19	US-10-437-963-67845	Sequence 67845, A
25	43.2	3.1	2110	19	US-10-437-963-10628	Sequence 10628, A
26	43	3.1	1305	17	US-10-282-122A-11478	Sequence 11478, A
27	43	3.1	1818	17	US-10-369-493-27231	Sequence 27231, A
28	43	3.1	2021	19	US-10-437-963-49750	Sequence 49750, A
29	42.8	3.1	956	18	US-10-425-114-24149	Sequence 24149, A
30	42.8	3.1	984	20	US-10-425-115-118414	Sequence 118414, A
31	42.8	3.1	993	19	US-10-437-963-93714	Sequence 93714, A
32	42.6	3.1	2213	21	US-10-772-636-27	Sequence 27, Appl
33	42.6	3.1	2231	17	US-10-282-122A-15175	Sequence 15175, A
34	42.4	3.0	1122	15	US-10-156-761-4985	Sequence 4985, Ap
35	42.4	3.0	1305	17	US-10-282-122A-12878	Sequence 12878, A
36	42.4	3.0	1470	9	US-09-815-242-7855	Sequence 7855, Ap
37	42.4	3.0	9025608	15	US-10-156-761-1	Sequence 1, Appli
38	42.2	3.0	375	18	US-10-425-114-4691	Sequence 4691, Ap
39	42.2	3.0	501	18	US-10-425-114-319	Sequence 319, Appl
40	42.2	3.0	726	18	US-10-425-114-58	Sequence 58, Appl
41	42.2	3.0	1525	17	US-10-369-493-27882	Sequence 27882, A
42	42.2	3.0	1866	20	US-10-425-115-74324	Sequence 74324, A
43	42	3.0	536	17	US-10-338-110-119	Sequence 119, App
44	41.6	3.0	900	17	US-10-369-493-41628	Sequence 41628, A
45	41.6	3.0	1271	19	US-10-437-963-38282	Sequence 38282, A

ALIGNMENTS

```
RESULT 1
US-10-689-200-1
Sequence 1, Application US/10689200
Publication No. US20040126848A1
GENERAL INFORMATION:
APPLICANT: Diccionario, Deana J.
APPLICANT: Ni, Hao
APPLICANT: Ye, Rick
APPLICANT: Picataggio, Stephen
APPLICANT: Wang, Tao
APPLICANT: Seip, John E.
TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN CL METABOLIZING BACTERIA
FILE REFERENCE: CL1747 US NA
CURRENT FILING DATE: 2003-10-20
PRIOR APPLICATION NUMBER: 60/419,872
PRIOR FILING DATE: 2002-10-21
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1392
TYPE: DNA
ORGANISM: Methylobionas sp. 16a
US-10-689-200-1
Query Match
Best Local Similarity 100.0%; Score 1392; DB 19; Length 1392;
Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 ATGAATACCATCTTAGATCGAGCTCGAAGAAATTTATTATGACGTATCGGCTTCGCTA 60
```

```

Db      1 ATGAAAACCATATGATGATGAGCTCGAAGAAATGTATATGACGTTATCGGCTTGCTA 60
Qy      61 GCGGTTGGGGTTTGACGATGCGCCGAGTGCAGGCGGCGATTGGGCAAGCTGAGAAAAGAA 120
Db      61 GCGGTTGGGGTTTGACGATGCGCCGAGTGCAGGCGGCGATTGGGCAAGCTGAGAAAAGAA 120
Qy      121 GATTGAAATTCGGCTTCATCAAGCTCACCGACATGGCGCCGCTGGCGGGTGGCCGCGAA 180
Db      121 GATTGAAATTCGGCTTCATCAAGCTCACCGACATGGCGCCGCTGGCGGGTGGCCGCGAA 180
Qy      181 AAAGGCTTCTTGAGAGACGAGGCGCTGTTGTCGAACTGGAAGCCGACGCGCAACTGGAAG 240
Db      181 AAAGGCTTCTTGAGAGACGAGGCGCTGTTGTCGAACTGGAAGCCGACGCGCAACTGGAAG 240
Qy      241 GGGGTATGATGATGAGGTGCGTGAATGGCGAATGGCGGCTCGACATGCTGGGCGCCGAG 300
Db      241 GGGGTATGATGATGAGGTGCGTGAATGGCGAATGGCGGCTCGACATGCTGGGCGCCGAG 300
Qy      301 CCGTTAGCGGCGACGCTTGCTTCGCGACCAAGGCGGATTCGAGGTGCGCTTCAGCATG 360
Db      301 CCGTTAGCGGCGACGCTTGCTTCGCGACCAAGGCGGATTCGAGGTGCGCTTCAGCATG 360
Qy      361 GCGTTCAACGGCAACCGATTCAGGTGTCGATGAATTCGGCATCAGATGAAGCCGAAC 420
Db      361 GCGTTCAACGGCAACCGATTCAGGTGTCGATGAATTCGGCATCAGATGAAGCCGAAC 420
Qy      421 ATACGCGTGAAGGCGGCTTAAACCGGTGACATCCGATCAAGGCGGATTCGATGAAGCCGAC 480
Db      421 ATACGCGTGAAGGCGGCTTAAACCGGTGACATCCGATCAAGGCGGATTCGATGAAGCCGAC 480
Qy      481 GTGAAAAAATACAAAGCCGAGGCAAGCGGTTCAATATGAGCGATGACCTTCCGCGCGGA 540
Db      481 GTGAAAAAATACAAAGCCGAGGCAAGCGGTTCAATATGAGCGATGACCTTCCGCGCGGA 540
Qy      541 TCTCACACATCAAACTGCGTTACTGCTGCGCGGCGGTATCAATCTGCTACTAT 600
Db      541 TCTCACACATCAAACTGCGTTACTGCTGCGCGGCGGTATCAATCTGCTACTAT 600
Qy      601 TGGCGCGCGACGAGCACTTCCGCGCAATCCGCGGAGAGCGGCTGTTGTCGCGTACCCG 660
Db      601 TGGCGCGCGACGAGCACTTCCGCGCAATCCGCGGAGAGCGGCTGTTGTCGCGTACCCG 660
Qy      661 CCGCGCGCAATGCGCTGCAAGCGTGAAGCGCGACCATTTTGGTTATTCGCTCGCGAG 720
Db      661 CCGCGCGCAATGCGCTGCAAGCGTGAAGCGCGACCATTTTGGTTATTCGCTCGCGAG 720
Qy      721 CCGTGAACCAAGCGCGCTGTTCAAGGCGATGCGCGTGCCTGATCAACCGATGAAGAA 780
Db      721 CCGTGAACCAAGCGCGCTGTTCAAGGCGATGCGCGTGCCTGATCAACCGATGAAGAA 780
Qy      781 CTCTGGAAGACACGCGCGGAAAAAGTCTTCGCGGTGACCAAACAATGGCGGAGAAATAC 840
Db      781 CTCTGGAAGACACGCGCGGAAAAAGTCTTCGCGGTGACCAAACAATGGCGGAGAAATAC 840
Qy      841 CCCAACACCTATCTGGCGGTGACCAAGCGCTGATTCGGGCGCGCATCTGCTGAGCGCC 900
Db      841 CCCAACACCTATCTGGCGGTGACCAAGCGCTGATTCGGGCGCGCATCTGCTGAGCGCC 900
Qy      901 GACATATACAAAGACCGCAAGGAGCATTCGAAATGCTGGCGCAAAAACAATACGTCGT 960
Db      901 GACATATACAAAGACCGCAAGGAGCATTCGAAATGCTGGCGCAAAAACAATACGTCGT 960
Qy      961 GGTGAAGTGAAGTGTGCGCGCTGACGAAGACGCGCATTTGGAAATGACAAAGAGCAT 1020
Db      961 GGTGAAGTGAAGTGTGCGCGCTGACGAAGACGCGCATTTGGAAATGACAAAGAGCAT 1020
Qy      1021 AAACGCGCGCTACCGGACTTCAACCTTCTTTCGCGACGCGCGCAGCTATCCGCTCTAC 1080
Db      1021 AAACGCGCGCTACCGGACTTCAACCTTCTTTCGCGACGCGCGCAGCTATCCGCTCTAC 1080
Qy      1081 AGCAGTGCAGTCTGATCTGACCCACGCTGAAGCGCTGGGCGCATGATCAATGAATTCAA 1140

```

```

Db      1081 AGCAGTGCAGTCTGATCTGACCCAGCTGAGCGCGTGGCGCATGATCAATTCAA 1140
Qy      1141 CCGGCAACCTGATCTGATCTGATCTGCGCGCAAGACGTCACCGCGGAGCATCTATCGCC 1200
Db      1141 CCGGCAACCTGATCTGATCTGATCTGCGCGCAAGACGTCACCGCGGAGCATCTATCGCC 1200
Qy      1201 GCGGCGCAAGGAACTGGTTCGACAGAGGCAAGGCGCAAGGCAAGCTTCCTGCGCATAC 1260
Db      1201 GCGGCGCAAGGAACTGGTTCGACAGAGGCAAGGCGCAAGGCAAGCTTCCTGCGCATAC 1260
Qy      1261 TCGATCAAGCGCTGCGCAAAATTTCTTCAATCGACAAAGTGCCTGATGCGCAAGCC 1320
Db      1261 TCGATCAAGCGCTGCGCAAAATTTCTTCAATCGACAAAGTGCCTGATGCGCAAGCC 1320
Qy      1321 AACGATTAACCTGCGCAAGTTTTCGATGAGTTCGAAAGGCAAGCAACCGTAGCCGCGC 1380
Db      1321 AACGATTAACCTGCGCAAGTTTTCGATGAGTTCGAAAGGCAAGCAACCGTAGCCGCGC 1380
Qy      1381 AAGGTCGTGAT 1392
Db      1381 AAGGTCGTGAT 1392

```

```

RESULT 2
US-10-689-200-3
; Sequence 3, Application US/10689200
; Publication No. US20040126848A1
; GENERAL INFORMATION:
; APPLICANT: Dicosimo, Deana J.
; APPLICANT: Ni, Hao
; APPLICANT: Ye, Rick
; APPLICANT: Picataggio, Stephen
; APPLICANT: Wang, Tao
; APPLICANT: Seip, John E.
; TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN CL METABOLIZING BACTERIA
; FILE REFERENCE: CL1747 US NA
; CURRENT APPLICATION NUMBER: US/10/689,200
; PRIOR FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 60/419,872
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 3
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Methylobionas sp. 16a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (409)..(411)
; OTHER INFORMATION: ATG start site of nra gene
US-10-689-200-3

```

```

Query Match      10.3%; Score 143; DB 19; Length 551;
Best Local Similarity 100.0%; Pred. No. 7.7e-34; Indels 0; Gaps 0;
Matches 143; Conservative 0; Mismatches 0;

```

```

Qy      1 ATGAAAACCATATGATGATGAGCTCGAAGAAATGTATATGACGTTATCGGCTTGCTA 60
Db      409 ATGAAAACCATATGATGATGAGCTCGAAGAAATGTATATGACGTTATCGGCTTGCTA 468
Qy      61 GCGGTTGGGGTTTGACGATGCGCCGAGTGCAGGCGGCGATTGGGCAAGCTGAGAAAAGAA 120
Db      469 GCGGTTGGGGTTTGACGATGCGCCGAGTGCAGGCGGCGATTGGGCAAGCTGAGAAAAGAA 528
Qy      121 GATTGAAATTCGGCTTCATCA 143
Db      529 GATTGAAATTCGGCTTCATCA 551

```

```

RESULT 3
US-10-282-122A-33791
; Sequence 33791, Application US/10282122A
; Publication No. US20040029129A1

```

```
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ LENGTH: 1209
/ TYPE: DNA
/ ORGANISM: Pseudomonas syringae
US-10-282-122A-33791

Query Match          5.2%; Score 71.8; DB 17; Length 1209;
Best Local Similarity 53.8%; Pred. No. 2.1e-11;
Matches 148; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 102 TGGCAAGCTGAAAAGAGATTGAAATTCGGCTTCATCAAGCTCAGCGACATGGCGCC 161
DB 45 TGACGCCCGCGAATATGAACAGCTGACGTGGGCTTCATGGCACTGACGACCTGCGGCC 104
QY 162 GCTGGGGGTGGCGCGCGAAAAGGCTTCTTCGAGACGAGGGGCTGTTCGTGCAACTGGA 221
DB 105 GCTGGTGTGTGCTGCCACCCAGGGGTTTGCCAGCCTTACGGGCTGAGCCTGAACCTGAA 164
QY 222 AGCGCAGGCGCAACTGGAAGGTGTGATGATAGGAGTGTGATGCGAACTGAGCGGCTC 281
DB 165 AGCGCAGAGCTCTCTGGGCGCGGGCTGCGGTGACAGGCTGTGACGGGGAGCTTCAAGCGCG 224
QY 282 GCACATGCTGGCGCGCGCGCGCTTAAAGCGGCGCAAGGCTTGGCGCACCAAGGCCATAT 341
DB 225 ACATAGCTGTATGCTTATGCTATGCGCTGAGAGCTGGGCGATAGCGCGCGCGCGAGAC 284
QY 342 CGAGTGGCGGCTTCAAGCATGAGGCTTCAACGCGCAAGC 376
DB 285 CGACATGGCCATCTCATGGGCTCAACCAAGACG 319

RESULT 4
US-10-282-122A-11998
```

```
/ Sequence 11998, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ LENGTH: 1173
/ TYPE: DNA
/ ORGANISM: Burkholderia cepacia
US-10-282-122A-11998

Query Match          4.5%; Score 62.2; DB 17; Length 1173;
Best Local Similarity 58.4%; Pred. No. 2.1e-08;
Matches 128; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 131 TCGGCTTCACTCAAGCTCAGCGATGCGCGCGCTGGCGCGCGGAAAAGGCTTCT 190
DB 116 TCGGCTATATGCGGATCAGACGACGCGCGCGCGCTGTGTGCTGCCACACACAGGCTATT 175
QY 191 TCGAGGACGAGGCGCTGT--TCGTGCAACTGGAACGCGACCACTGGAAGGTGTGA 247
DB 176 TCGCGTCCGAAGGGCTGAGCGGTGAGCAACCGAAGCTGTGGTACTGGGCGCAAGCTGG 235
QY 248 TGAATAGGTTCTGGAATGCGGAAGTGAAGGCTTCGACATGCTGGGGCGCGCGCTTGA 307
DB 236 TCGAGGCGTTCTGTCCGCGCGAGGTCAACGTGTGCACTGCTGTGCGCGATGACGCTGT 295
QY 308 CGGCGCAGGCTTGGCTTGGCGCACCAAGCGCGATATCGAGG 346
DB 296 GGGCGGCTACGCGCAGCCGCGCGCGCGCGCGCAAGTGTGG 334

RESULT 5
US-10-282-122A-11492
/ Sequence 11492, Application US/10282122A
/ Publication No. US20040029129A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11492
/ LENGTH: 576
/ TYPE: DNA
/ ORGANISM: Burkholderia cepacia
US-10-282-122A-11492

Query Match          4.3%; Score 59.2; DB 17; Length 576;
Best Local Similarity 57.9%; Pred. No. 1.4e-07;
Matches 125; Conservative 0; Mismatches 88; Indels 3; Gaps 1;

QY 131 TCGGCTTCATCAAGCTCACCGACATGCGCGCTGGCGGTGGCGCGCAAAAAGGCTTCT 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 116 TCGGCTATCTGCCGATCACGACGCGCGCGCTGCTGTGCGCCACACACAGGCTATT 175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 TCGAGACGAGGCGCTGT--TCGTGCACTGGAAAGCGACGCAACTGGAAAGTGTGA 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 176 TCGGCTCCGAAGGCGTGAAGGTGAGGAAACCAAGCGTGGCGCACTGGCGCAACTG 235
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 248 TGGATAGGCTGTGGAATGGCGAACTGGAAGCGCTGCGCATGTGGCGCGCGCGCTTGA 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 236 TCGAGGCGTTCCGTCCGCGGACAGGTCAACGTCTGTGACCGTGTGGCGATGACGCTGT 295
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 CGGCGAGCGTTGGCTTGGGACGACGCAAGGCGGATATCG 343
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 296 GGGCAGCGTACGGCAGCGCGCGCGCGCGCAAGGTGG 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 6
US-10-282-122A-30186
/ Sequence 30186, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
```

```
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 30186
/ LENGTH: 1200
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-30186

Query Match          4.2%; Score 58; DB 17; Length 1200;
Best Local Similarity 56.2%; Pred. No. 4.5e-07;
Matches 153; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 131 TCGGCTTCATCAAGCTCACCGACATGCGCGCGCTGGCGCGCGCAAAAAGGCTTCT 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 137 TCGGCTACCTGCCGATCACCGACGCGCGCTTCTGTGCGCCATGCGCAACGCGCTGT 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 TCGAGACGAGGCGCTGTTCGTGCAACTGGAAAGCGCAGCC--AACTGGAAGTGTGA 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 197 TCGAGGCGGAGGCGCATTCAGAGCGCAACGCGCGGTGCTGTGCGTACTGGCGCAGTGA 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 248 TGGATAGGCTGTGGAATGGCGAACTGGAACGCGTCCGCAATGTGGCGCGCGCGCTTGA 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 257 TCGAGGCGCTTCATCTCCGCGCAGTCAAGTATCACCCTGTGTGCGCGGATGACCGTCT 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 CGGCGAGCGTTGGCTTGGCGCAACGAGCGGATATCGAGGTGCGCTTACAGCATGGGCTTCA 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 317 GGGGCC-----CGCTACGGCAGCAAGTGTCCGCGCAAGTGTGGCTTGGAAACACGTGC 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 368 ACGGCAAGCGGATTAAGGTGTTCAAATGAATC 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 371 GCGGCTCGGGGCTCACGGTGGCTCCGAGATC 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 7
US-10-389-647-149
/ Sequence 149, Application US/10389647
/ Publication No. US2004003549A1
```



```

PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/230,335
PRIORITY FILING DATE: 2000-09-06
PRIORITY APPLICATION NUMBER: 60/230,347
PRIORITY FILING DATE: 2000-09-09
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/267,636
PRIORITY FILING DATE: 2001-02-09
PRIORITY APPLICATION NUMBER: 60/269,108
PRIORITY FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM
NUMBER OF SEQ ID NOS: 76614
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 25468
LENGTH: 7656
TYPE: DNA

```

Query Match	3.4%	Score	47.8	DB	17	Length	7656
Best Local Similarity	54.2%	Pred. No.	0.0015				
Matches	97	Conservative	0	Mismatches	82	Indels	0
						Gaps	0

QY 159 GCGCGTGGCGGGTGGCGCGCCGAAAAGAGCCTTCTTGAGAGACGAGGGGCGCTGTTCGGCACT 218
 Db 1929 GACCATGTGTCTGGCGCGCGCGCTCGGGCGTCACCGCGGACGACGTCCGGGTGTGTCTGCA 1988
 QY 219 GGAAGCGCAGGCCCACTGGAAGGTGTGTATGATGATAGGGTCTGTAAATGGCGAACTGACCG 278
 Db 1989 GGGCGTGTGGACCGCACCGACCGATGTCTGTGGCGGTGTGGAGGACGACCGCGCGCG 2048
 QY 279 CTGCGCACATCTGTGGCGCGCGCGCGCTTAAAGCGGCGCAGCGGCTTGGCTTCGGCACCAAGCGCG 337
 Db 2049 CTGGAACCTGACGCGCGCGCGCGAGCGCGGCTCGATGCGCGCGCGCGACTGCTGTGACGTGG 2107

RESULT 14
 US-10-156--761.3146
 ; Sequence 3146, Application US/10156761
 ; Publication No. US20030119018A1

```

RESULT 14
US-10-156-761-3146
; Sequence 3146, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMOBA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3146
; LENGTH: 3048
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3048)
; US-10-156-761-3146

```

Query Match 3.3%; Score 46.4; DB 15; Length 3048;
Best Local Similarity 50.0%; Pred. No. 0.0028;
Matches 116; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Search completed: September 16, 2005, 05:01:18
Job time : 949 secs

QY 69 GGGTTGACGATGTCGCCGCGCATGTGCGGCGAGTTGGCAAGCTGGAAGAAAGAGATTGAA 128
DB 645 GGTGCTGACGACCGCCCTTGACGTCGTGTGCGGCCGCTACACCGCCAGAGAGACATCCC 704
QY 129 ATTGGGCTTCATCAAGGTCAACGACATGAGCGCGCTGGCGCTGCCCGGAAAAAGGCTT 188
DB 705 GCTCGGATCCCATGTGTCGGCCGACGAGCCCGAATCTGAGAGACGTGTCGCGCTGTT 764
QY 189 CTTGAGAGACGAGGCGCTGTTCTGTCGCACTGGAAGCGCAGCCCACTGGAAGGTGTGAT 248
DB 765 CATCAACATGCTGTGTGTGCGCACCGATCTGTCCGGGAGACCCACCTTCCGCGAGCTGAT 824
QY 249 GATAGGGTCTGTAATGCGCACTGACGCGCTGCGCACATGTCGCGCGCGCG 300
DB 825 CGAGCGGGTCCGCGACGCCAATGACGTGTACGACAAACGAGACGTGCGCG 876

RESULT 15

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.3%; Score 46.4; DB 15; Length 9025608;
Best Local Similarity 50.0%; Pred. No. 0.076;
Matches 116; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 69 GGGTTGACGATGTCGCCGCGCATGTGCGGCGAGTTGGCAAGCTGGAAGAAAGAGATTGAA 128
DB 3936846 GGTGCTGACGACCGCCCTTGACGTCGTGTGCGGCCGCTACACCGCCAGAGAGACATCCC 3936787
QY 129 ATTGGGCTTCATCAAGGTCAACGACATGAGCGCGCTGGCGCTGCCCGGAAAAAGGCTT 188
DB 3936786 GCTCGGATCCCATGTGTCGGCCGACGAGCCCGAATCTGAGAGAGTGTGCGCGCTGTT 3936727
QY 189 CTTGAGAGACGAGGCGCTGTTCTGTCGCACTGGAAGCGCAGCCCACTGGAAGGTGTGAT 248
DB 3936726 CATCAACATGCTGTGTGTGCGCACCGATCTGTCCGGGAGACCCACCTTCCGCGAGCTGAT 3936667
QY 249 GATAGGGTCTGTAATGCGCACTGACGCGCTGCGCACATGTCGCGCGCGCGCG 300
DB 3936666 CGAGCGGGTCCGCGACGCCAATGACGTGTACGACAAACGAGACGTGCGCG 3936615

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 00:58:27 ; Search time 42 Seconds

(without alignments)
1062.966 Million cell updates/sec

Title: US-10-689-200-2

Perfect score: 2443
Sequence: 1 MTTIRSSSKULLTSLASL.....AKPAIGKQIVAGKQVD 464

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651	26.6	670	2	S77387
2	627.5	25.7	430	2	AC3037
3	627.5	25.7	471	2	H98248
4	616.5	25.2	441	2	E95928
5	611.5	25.0	431	2	AD3037
6	611.5	25.0	431	2	G98248
7	610.5	25.0	440	2	AG1882
8	587	24.0	657	2	A11882
9	576	23.6	430	2	C95300
10	572	23.4	443	2	J02134
11	549.5	22.5	442	2	G87324
12	545.5	22.3	442	2	S58738
13	538.5	22.0	667	2	AH2165
14	534.5	21.9	667	2	S75959
15	525	21.5	446	2	S77389
16	502.5	20.6	458	2	AF2165
17	494	20.2	402	2	F83422
18	476	19.5	659	2	S30893
19	473.5	19.4	452	2	S75956
20	470.5	19.3	427	2	AF2222
21	444	18.2	450	2	A44751
22	398.5	16.3	392	2	S52248
23	390.5	16.0	426	2	AF2663
24	390.5	16.0	426	2	E97445
25	355	14.5	625	2	S58789
26	249	10.2	225	2	B55581
27	240	9.8	399	2	A83354
28	125	5.1	337	2	B75370
29	124.5	5.1	333	2	AB3080

30	124.5	5.1	333	2	G98206	hypothetical prote
31	116.5	4.8	468	2	A83359	hypothetical prote
32	114	4.7	325	2	H75414	probable ABC trans
33	114	4.7	336	2	A83801	ABC transporter (s
34	111	4.5	310	2	AC3238	hypothetical prote
35	109.5	4.5	300	2	H69260	thiamin biosynthes
36	106.5	4.4	675	2	D48492	kpac protein - Bac
37	106	4.3	483	2	G86811	sugar ABC transp
38	106	4.3	1273	2	T34558	hypothetical prote
39	105.5	4.3	339	2	E95962	probable taurine u
40	105.5	4.3	566	2	A40589	cellulase (EC 3.2.
41	104.5	4.3	1135	2	T14699	DNA polymerase III
42	104.5	4.3	1174	2	T15021	probable DNA poly
43	103	4.2	897	2	AC2001	hypothetical prote
44	102	4.2	301	2	B95947	phosphate uptake A
45	102	4.2	1291	2	I40631	non-proteolytic bo

ALIGNMENTS

RESULT 1

S77387
nitrate transport protein C-2 - *Synechocystis* sp. (strain PCC 6803)

N/Alternate names: protein s111452

C/Species: *Synechocystis* sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004

C/Accession: S77387

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

8.

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S77387

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-670 <KAN>

A/Cross-references: UNIPROT:P73450; EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BA0174;

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Genetic:

A/gene: nrtC-2

C/superfamily: ATP-binding cassette homology

C/Keywords: ATP; nucleotide binding; P-loop

F/25-315/Domain: ATP-binding cassette homology <ABC>

F/42-49/Region: nucleotide-binding motif A (P-loop)

Query Match 26.6%; Score 651; DB 2; Length 670;

Best Local Similarity 36.3%; Pred. No. 9.7e-40;

Matches 158; Conservative 68; Mismatches 153; Indels 56; Gaps 13;

QY	28	ADPVGAVGK-----LEKEDLKFGFKLTDMAPLAAVAKGFEEDEGL-FVQLEAQA 77
DB	260	AKKVGAVSQFAPAMGNGKELINDGRIPIIDCAPLVAAKKGFQKRGLEQVNLVXEP 319
QY	78	NKKVVMDRVNVGSELDSHMLAPLAASVGFSTKADIEV--PFSKGFNGNAITVSNEIWH 135
DB	320	SMOAIADGIRREKRLDGAQVAVAGMPPLATLTGMGSKTPIPLVATVAMVMSRNAGATLTSKFAEP 379
QY	136	QMKPNIPLEGKGFVPIKADYLPVVEKTKAGSKFPNNAMTTPAGSHNFKLAVYMLAAGCI 195
DB	380	-----AGVTLDELR-----LKLAETPDQSTLGMVHPASMNLLRYWLASGSI 424
QY	196	NGGVYSPQDIDSGQIGADALLSVTPPQPMSTLEAGTTFGVCVGEPMNQOAVFKGIGVVP 255
DB	425	D-----EPDDIN-----LMRLPPQMVNSLEAGNIDGFCVGEPMNSYAAKQNLGYVI 471
QY	256	ITDBELMKOTPEKPVGVTKQWAEKTPNTYLAATKALIRAAIWLADNNKRRKAEATEMLAQ 315
DB	472	ATVDLIDWNGHPRKPVGLMREEWNVKKYPATIALVKKALLEACEY--CDDRHRROEIIDYLA 529
QY	316	KQVGVADVLEVLAAASNNGTREYEKD-DKRALPFPNFFFRHGAYPSSASVWVLTQLRRNG 374

QY 211 GADALLSVTPPQMPSTLEAGTIFGYCVGEPMNOQAVFKIGIVPITDELMKOTPEKYE 270
DB 191 --DYKLVVPPPTSDLAAGALDGFVGAPMNIVAERVGRIVAAKODLMSABEKYI 248
QY 271 GVTQMAEKYPNTYLAATKALIRAIWLD-ADNNKRNKEAIEMLAQOYGVADVEVLAAS 329
DB 249 GMPPEWMSQOETVGRLLTALDAAASWCDLADNHDLSGA---LADPRYIGAPQSIIRRV 305
QY 330 MNGTFERE-KDDGRALPDPTFFRHGASYPSSAVVYLTQLRRKMINEFKPDNMTLDT 388
DB 306 LAGEFSDSQGNRRVIEKYTFPHGDHANYPQSQSLMYSQMRWG---QAELESTGVNA 362
QY 389 AKVVPDYLAAKELVAEKAEDFPADTSIK---PSQNFIDKVPFDANKENDYI 444
DB 363 ALSAYRPDIYRAA-----LGDGKA-----PDADIRIEGDEGRFVDFGVFDPDADIAGYV 413
QY 445 AKFAI 449
DB 414 NSFAY 418

RESULT 7

AG1882
nitrate transport nitrate-binding protein ntrA [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AG1882
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yaeuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AG1882
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-440 <KUR>
A/Cross-references: UNIPROT:Q44292; GB:BA000019; PIDN:BA072566.1; PID:g717129954; GSPDB:G
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: ntrA

Query Match 25.0%; Score 610.5; DB 2; Length 440;
Best Local Similarity 33.4%; Pred. No. 5e-37;
Matches 158; Conservative 71; Mismatches 173; Indels 71; Gaps 17;
QY 8 SSKLLLT---LSASLAWGLTI-----APDVGAVG-KLEKEDLKF 44
DB 5 SRRKFLPTTGAAGAAASLTVHGTSSNGSQAATGEOAPSAAPANVSAANAPEKVEITKAL 64
QY 45 GPKLTMDPLAAVAEKGFEDEGLF-VOLEAQANKVWDRY---VNGELDGHMLAP 99
DB 65 GFLPLDADPLIAKEKGFPKAKYGMWDIEVTKQKSPVTRDKLTGSSGGDIGHAHTLSB 124
QY 100 APLAASVGFCTKADIEVPSMGFNAGATVSNIEIWHQMKENIPLEGKRPVHPKADYLP 159
DB 125 MPTLMTI--NDKVPMVILARLNTNGQALISAEKF---KELNVLE-----SKSLMD 170
QY 160 VVEKYRAEKGFPNMAMTFPAGSHNIKRWTYLAAGINPGYSPPODISQIGADLLSTV 219
DB 171 AATKAKADKALMGITPFGTHDLMRWYLAAGINP-----DQDVLEAV 217
QY 220 PPPQMPSTLEAGTIFGYCVGEPMNOQAVFKIGIVPITDELMKOTPEKVEITKQMAEK 279
DB 218 PPPQMPANMKVNTVDGCVGEPMNAQVLNOKIGYSALVTGELMKDPEKAFSRQWIMQ 277
QY 280 YPNTYLAATKALIRAIWLDADNNKRNKEAIEMLAQOYGVADVEVLAASMNGTFEYED 339
DB 278 NPMAAQALIMAILLEAQOM--CDKAEKKEBCKICSRKTFNVAAADIIIRAKENIDY--G 333
QY 340 DKRALPDF--NTFFRHGASYPSSAVVYLTQLRRKMINEFKPDNMTYLTAKNMYRPD 396

DB 334 DGRKEQNFARHMKFWADNASYPYKSHDIFLIEDIRMGYLPK---DTKQDILVNGVNKED 390
QY 397 IYLAAKEIVAGKAKAEDFPADTSIKSQNFIDKVPDANKPDYLAKEFI 449
DB 391 LMKKAQ--ATGVADAE-IPASS-RGVETFF-DGVKFPDPKPEPEYNSLKI 437

RESULT 8

AG1882
nitrate transport ATP-binding protein ntrC [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AG1882
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yaeuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AG1882
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-657 <KUR>
A/Cross-references: UNIPROT:Q8Y276; GB:BA000019; PIDN:BA072568.1; PID:g717129956; GSPDB:G
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: ntrC
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 24.0%; Score 587; DB 2; Length 657;
Best Local Similarity 35.8%; Pred. No. 4.6e-35;
Matches 150; Conservative 74; Mismatches 139; Indels 56; Gaps 17;

QY 39 KEDLRGPKLTMDPLAAVAEKGFEDEGLF-VOLEAQANKVWDRYVNGELDGHML 97
DB 276 KAVIEIGFMPITDSAPLIVAEKGFPKAYGLDNVILNRANMQALATGVTGKDLAAQMV 335
QY 98 AAPLAASVGFCTKADIEV--PFSMGFNAGATVSNIEIWHQMKENIPLEGKRPVHPKAD 155
DB 336 ACPMPLALTIGASQPTPYVINALNLSRANALTFEKRILYNGVRSL-----AD 383
QY 156 YLKPVVEKYRAEKGFPNMAMTFPAGSHNIKRWTYLAAGINPGYSPPODISQIGADAL 215
DB 384 LKAVID--SSPDQLLTGTVHSAQMQLILRYWLAAGID-----PBDVG----- 427
QY 216 LSVTPPQMPSTLEAGTIFGYCVGEPMNOQAVFKIGIVPITDELMKOTPEKVEITKQ 275
DB 428 LTVIPPTQWVSQLKAGNIDGYCAGEPMNYQAVHDDLGVAAATALISWGGQPKVILGVRED 487
QY 276 WAKEYPNTYLAATKALIRAIWLDADNNKRNKEAIEMLAQOYGVADVEVLAASMNGTF- 334
DB 488 WAQKPEPYLINLVKALIBACK--CDDLARNEBELICREPY-----LDVNPAYVRSGL 541
QY 335 -EYKED--KRALPDPTFFRHGASYPSSAVVYLTQLRRKMINEFKPDNMTYLTAK 390
DB 542 DRYDGDGTPPOLTAANYQFYLNTKNTYDNRTEILMTIQMRWG-LTFP-PGMV-VEITE 598
QY 391 NVYRPDIYLAKEIVAGKAKAEDFPADTSIKSQNFIDKVPDANKPDYLAKEFI 449
DB 599 RVCRTDIFGAARDL-----GLLDIGEDDI--HLFDGL-FNPSBPILYLSLEI 646

RESULT 9

C95300
NtrA-type periplasmic nitrate transport binding protein, probable Sma0585 [imported] - S
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: C95300
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A/Reference number: A95262; MUID:21396509; PMID:11481432

[illegible][illegible]

Best Local Similarity 32.7%; Pred. No. 1,4e-32;
Matches 148; Conservative 68; Mismatches 179; Indels 57; Gaps 16;

QY 6 RSSSKLLTLTSLANM-GLTTPDVGAAGKLEKEDEKGFITLDTMAPLAAAEKGF 64
DB 16 RSLAMGAAATLAAKAKAFPGAAHAGAP---EVAKARIGFTALTDSSPLIAKRGLE 71
QY 65 EDEGL-FVQLEAQAQNMKVMDRVY---NGELDGSMLAPAPLAASVGF---GTKADIEV 116
DB 72 AKKGMPDIEVYKQASMAATRDNLVLGAERGIQGAHILTMPLMTGTTTGGAFFPMYI 131
QY 117 PSMGFNGNAITVSNELIWHQMKPNIPLEGKPVHPIKADYLKPVKEKYAKGKPFMMAMT 176
DB 132 LALNTNGQGISVGNDL-KSVKVLNSAGA-----KAKFL-----QMKAGNIAYKAMT 179
QY 177 FPAQSHNIKLRMYLAAAGINPGYVSPPODISQIGADALLSTYPPQMPSTLEAGTTFG 236
DB 180 FPGTIDHLMTRYMLAAGINP-----DVDVSTIVIPPPQVAMNMAAGTQDAF 226
QY 237 CVGEPMNQAVFAGIGVPIITDEELMKDTEPEKVFQVTKQMAEKYPNTYLAATKALIRAAI 296
DB 227 CVGEPMNGQVLNKGVTACTLTSELMMNHPEKALGMRASVWDKYPRAAQITAAVQEAQM 286
QY 297 WLDADNNKRKEAIEMLAQOYVGADVEVLAASNGTFEYKDKRAL---PDPTFFPR 353
DB 287 W--CDKXANLPQMCISYSGRQYVNPVWGDIPLRIGTVDY--GDGRTLKNSPRMKFWAD 342
QY 354 GASYPSSAVWYLTQLRKGMINEFKPDWYLDTKANVRRPDIYLAALAEVLAEGKAKA 413
DB 343 NASPFKSHDLWFLTEBIRWGLVPQKTNTKALVD--KVRSDIWRALAAKSIGQSG--- 395

QY 414 EDPADTSIRKSNFPIDKVPDPANKRNDYLA 445
DB 396 ---PAGDS-RGVERFFDGKV-FDPANFGAYLA 422

RESULT 12
S58738
nitrate-binding protein ntrA precursor, periplasmic [similarity] - Phormidium lamiosum
C/Species: Phormidium lamiosum
C/Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C/Accession: S58738; S56641; S62124
R/Merchan, F.; Kindle, K.L.; llama, M.J.; Serra, J.L.; Fernandez, E.
Plant Mol. Biol. 28, 759-766, 1995
A/Title: Cloning and sequencing of the nitrate transport system from the thermophilic, *Phormidium lamiosum*
Cus sp. PCC 7942
A/Reference number: S58738; MUID:95375238; PMID:7647306
A/Accession: S58738
A/Molecule type: DNA
A/Residues: 1-442 <MER>
A/Cross-references: UNIPROT:Q51880; EMBL:Z19598; NID:91154890; PID:CAW79656.1; PID:e218
R/Merchan, F.; Pileto, R.; Kindle, K.L.; llama, M.J.; Serra, J.L.; Fernandez, E.
Plant Mol. Biol. 27, 1037-1042, 1995
A/Title: Isolation, sequence and expression in *Escherichia coli* of the nitrite reductase
A/Reference number: S56640; MUID:95284340; PMID:7766873
A/Accession: S56641
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-67, 'A', 69-442 <MEM>
A/Cross-references: EMBL:Z19538
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C/Genetics:
A/Gene: ntrA
C/Keywords: nitrate transport; periplasmic space

Query Match 22.3%; Score 545.5; DB 2; Length 442;
Best Local Similarity 33.3%; Pred. No. 2.9e-32;
Matches 145; Conservative 67; Mismatches 175; Indels 49; Gaps 14;

QY 27 IAP-DVGAAGKLEKEDLKGFITLDTMAPLAAAEKGFEDBGL-FVQLEAQAQNMKVMD 84
DB 44 LAPAVAAADTPEVTTAKGFCPPDGLAFLIAKRGLEKAFKQMPDVEVLQKASMAVTRD 103

QY 85 RVV---NGELDGSMLAPAPLAASVGFCTKADIEVPS---MGFNAGAITVSNELIWHQ 136
DB 104 NIELSGGGGIDGAIHILTMPLMYSAGTITKNQVNPANNIILARLVNNGGICLAATY--- 160
QY 137 MKPNIPLEGKRPVHPIKADYLKPVVEKYAKGKPFMMAMTPPAGSHNIKLRMYLAAAGIN 196
DB 161 ---KDLKIGTSSPLKEAFAKA-----KAEKGEIKAAVTFFGGTIDLMRWYLSAGCID 211
QY 197 PGYSPPODISQIGADALLSTYPPQMPSTLEAGTTFGCVGEPMNQAVFKGIGVPI 256
DB 212 ---PDKDIS-----TIVPPQVAMNKNMMEFPCVGEPMNPAQTVOGLGYNM 258
QY 257 TDEELMKDTEPEKVFQVTKQMAEKYPNTYLAATKALIRAAIWLADNNKRKEAIEMLAQ 316
DB 259 TTGELMKHPEKAFAMRADWVEGHHKAKALLMAVQEAQIW--CDDPANKEMCGVSGR 316
QY 317 QYVGADVEVLAASNGTFEYKDKRALPDPTFFRHGASYSYSASAWYLTQLRKGM 376
DB 317 EWFKVPVEDILERSKGNFDLGVQLENSFLMKFMNDAASYPFKSHDLFTEBIRWGLY 376
QY 377 NEFKDNNWLDPAKNVYRPDIYLAAKELVNEGKAKAEDFPADTSIKSONFPIDKVPRD 436
DB 377 PADTDTKALVDA---VNREDLMREAAKAIQGEAA-----TPASTS-RGVETFF-DGVKPD 426
QY 437 ANKPNDYLAFAIGLK 452
DB 427 PENPSAYLS--ALXKX 440

RESULT 13
AH2165
bicarbonate transport ATP-binding protein cmcC [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AH2165
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saesomoto, S.; Matanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AH2165
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-667 <KUR>
A/Cross-references: UNIPROT:Q8Y747; GB:BA000019; PID:BAW74578.1; PID:g17131973; GSPDB:G
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: cmcC
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 22.0%; Score 538.5; DB 2; Length 667;
Best Local Similarity 33.3%; Pred. No. 1.7e-31;
Matches 140; Conservative 70; Mismatches 158; Indels 53; Gaps 14;

QY 30 DVGAAGKLEKEDLKGFITLDTMAPLAAAEKGFEDBGL-FVQLEAQAQNMKVMDRVYN 88
DB 270 DVARRG-LEKVNLEIGFLPLTACAPLAAAEKGFETTGKGLDENVLVRETSWKGIDGMKG 328
QY 89 GELDGSMLAPAPLAASVGFCTKADIEV--PSMGFNAGAITVSNELIWHQMKPNIPLEGK 146
DB 329 GYIDAQMPGMPMMLTIGGHDNQLPVYALTMTNRNGAIFLARFYD-----EGV 380
QY 147 KPVHPIKADYLKPVVEKYAKGKPFMMAMTPPAGSHNIKLRMYLAAAGINPGYSPPODI 206
DB 381 RSLSDPK-NYL-----LRTDQRIHMGVNHASMNILLRKWLAAAGIDP----- 424
QY 207 SQGIDADALLSTYPPQMPSTLEAGTTFGCVGEPMNQAVFAGIGVPIITDEELMKDTP 266
DB 425 ---DLDVDMKRTTPPAQWADLQNSIDSGYCVGEPMNRAAVENGTITATDLEWLGHP 480
QY 267 EKVFGVTKQMAEKYPNTYLAATKALIRAAIWLADNNKRKEAIEMLAQOYVGADVEVL 326

```

Db 481 GKVLGVREDMAERYPNTHIALTKALLACEY--CSRPENVEVARIIVAGRDVSTDDIYI 538
Qy 327 AASNGAFPEYEKDKRALPDP--NTFRRHGA-STPSYSSAAVWYLTQRRMGMINFEKPDN 383
Db 539 QLEDPNSLVCIDID--HPLRDYAHQHQFAESALNRPSTEQIWMISQALRMG--DTPPRN 594
Qy 384 WYLDATAKNVVPDIYLAAXELVAEGKAKDEFPADSIKRSQNFIDKVPEDANKPNDY 443
Db 595 W-VEVVERVCRVRFVSTAREL-----GLDLSYTRQPIELFDGTPFNADDPDIAY 642
Qy 444 L 444
Db 643 L 643

RESULT 14
nitrate transport protein C-1 - Synechocystis sp. (strain PCC 6803)
S75959
N:Alternate names: protein slr0043
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004
C:Accession: S75959
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
5.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75959
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-667 <KAN>
A:Cross-references: UNIPROT:Q55462; EMBL:D64006; GB:AB001339; NID:G1001291; PIDN:BA01080
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Gene: ntrC-1
C:Superfamily: ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:25-215/Domain: ATP-binding cassette homology <ABC>
F:42-49/Region: nucleotide-binding motif A (P-loop)

Query Match 21.9%; Score 534.5; DB 2; Length 667;
Best Local Similarity 31.7%; Pred. No. 3.3e-31;
Matches 146; Conservative 73; Mismatches 154; Indels 87; Gaps 16;

Qy 10 KLLLTLSASIAVWGLTIAPDVGAVGKLEKEDLKFGFIKLTDMAPLVAAEKGFEEDEGL 69
Db 261 KKLRAKKTITAIARHG-----LEKYNLELGIYVPLVACPLVAOKEGFFFAAGL 308
Qy 70 -FVQLEAQAAMKVMRVNNGELDGSIMLAAPLAASVGGTKADIEV--PFSMGFNGNA 126
Db 309 DEVSIVLETSMRGIVDGIAGYLDGAQMPAGMPTWMLAAGVREGSIIVSALMTTRNGA 368
Qy 127 IYTSNEMTQMKNPIPLEGGKPNPIADYIKPVVEKYKABGKPFNMAFTFPAGSHNIK 186
Db 369 ITTSKTKLYD-----GIYTAI-DFRQLL--ASDGRHTLTGMVHPSSMHLIL 413
Qy 187 RYMLAAGINPGYSPPODISGQIGADALLSVTPPMPSTLEAGTIFGYCVGEPNQA 246
Db 414 RYMLAANNIP-----DRVHLKTIIPPAQVADLKAGITIDGICVSEPMILRA 460
Qy 247 VFKGIVGVITDELMKDTPEKVGVTQKQAEKYPNTYLAVTKLRAALIMLADANNKNR 306
Db 461 SMEGAGSIATDLEIMQNHFGKVLGVREDMAIAHPNTHIALVYKALLBAAYC-ADPN-HE 518
Qy 307 KEAIEMLAKQYVADVEVL-----AASMGTFEYEDDKRALPDNTPFRHGAISY 358
Db 519 MEIRELLATROYLSTNIDYIHLGDPGRCRLGNPVEY-----SHHLFGDQPNRP 569
Qy 359 SYSASAVVYLTQLRWGMINEFKPDNWTLDATAKNVVPDIYLAAXEL-----V 406
Db 570 SRTEHLMMTQMARKWDIP--FRANW-VEILERVCRVGVSTARELGYVNTQROPAL 626

```

```

Qy 407 AEGKAAEDFPADTSIK-----PSQNFIDKVPEDANKP 440
Db 627 FDGKV-----FNADDPALVNLQTVIHRNFTIAEVLNPTP 662

RESULT 15
nitrate transport protein A-2 - Synechocystis sp. (strain PCC 6803)
S77389
N:Alternate names: nitrate transport 45K protein; protein sl11450
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S77389
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
5.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77389
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-446 <KAN>
A:Cross-references: UNIPROT:P73452; EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BA01174
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Gene: ntrA-2

Query Match 21.5%; Score 525; DB 2; Length 446;
Best Local Similarity 28.5%; Pred. No. 9.3e-31;
Matches 139; Conservative 78; Mismatches 186; Indels 84; Gaps 15;

```

```

Qy 1 MKTIIRSSKLLTTSASIAVWGL-----TIAPDVGAVGK 37
Db 1 MSNFSRSTRKRMFTAGAA-AIGGVVILHGTSTTTSTGTGSSSTDQALSPVLEGNA 59
Qy 38 EKEDLKFGFIKLTDMAPLVAAEKGFEEDEGL-FVQLEAQAAMKVMRVN-----NGELD 92
Db 60 EVTTAKLGIATLDAAPLIIAKEGKYAKGMDEVLKQASWGTTRDNLVLGASAGID 119
Qy 93 GSHMLAPPLAASVGF---GTRADIEVPSMGNGNAIYTSNEMTQMKNPIPLEGGKPV 149
Db 120 GAHILTPMPLLTMTGVTGDKPTPMYIILARLNVGGIDLGNNY-----KDLKVGTD 172
Qy 150 HPIKADYIKPVVEKYKABGKPFNMAFTFPAGSHNIKLRVYLAAGINPGYSPPODISG 209
Db 173 APIKAEPKATYDTPK-----VAMTFPGGTHDMIRYMLAAGNEP-----211
Qy 210 IGADALLSVTPPQMPSTLEAGTIFGYCVGEPNQAQVFGKIGVPIITDELMKDTPEKV 269
Db 212 -GRDFTIVPPAQVAVNVAAMESFCVGEPMPLGTVNOGVQYALTTGQLMKDHPKA 270
Qy 270 FGVTQKQAEKYPNTYLAVTKLRAALIMLADANNKNRKEAIEMLAKQYVADVEVLAAS 329
Db 271 FGRADWVQNPQKAAKLMAVMAEQM--CDQAEKKEECQITISGKEWFKYPFEDIIDR 328
Qy 330 MNGTFEYK-----DDKRALPDNTPFRHGAISYSSAAVWYLTQRRMGMINFEKPDN 384
Db 329 SKGIYVFGNGQETFEQDEIM--QKTVWDASTPIYSHQWFLTEIRNGYLPASDTYKA 385
Qy 385 YLDATAKNVVPDIYLAAXELVAEGKAKDEFPADTSIKRSQNF--FIDKVPEDANKPN 442
Db 386 IVD---KNVREDLMREAQAL-----EVPAQDIPS-----SPSRGIEFTFDGTFDPENPQA 434
Qy 443 YLAKEAI 449
Db 435 YLDSLKI 441

```

Search completed: September 16, 2005, 01:09:58
Job time : 46 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 00:49:17 ; Search time 174 Seconds
(without alignments)
1365.544 Million cell updates/sec

Title: US-10-689-200-2
Perfect score: 2443
Sequence: 1 MKTIISSSKKLLTSLASL.....AKPAIGLKQYVAGKRVVD 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt_03:*
1: uniprot_sprot:*
2: uniprot_tramb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1597	65.4	462	2	08KY55 azospirillum
2	1586	64.9	475	2	07MDT5
3	1572	64.3	487	2	08D5X8
4	1564	64.0	454	2	087HA7
5	1257.5	51.5	553	2	07UYV1
6	651	26.6	670	1	NRTC SYN3
7	627.5	25.7	430	2	08U925
8	627.5	25.7	471	2	07CTU1
9	616.5	25.2	441	2	092VK8
10	611.5	25.0	431	2	08U924
11	610.5	25.0	440	1	NRTA ANASP
12	610	25.0	429	2	07NEC4
13	594	24.3	437	2	07NKB1
14	587	24.0	657	2	08Y276
15	585.5	24.0	420	2	08Y2F5
16	582.5	23.8	439	2	089RG9
17	578.5	23.7	412	2	098H14
18	578	23.6	420	2	06D2V0
19	577	23.6	439	2	08DJ78
20	576	23.6	430	2	092Z21
21	575.5	23.6	418	2	048466
22	572	23.4	443	1	NRTA SYN7
23	567	23.2	654	2	08DJ76
24	566	23.2	425	2	098H12
25	559.5	22.9	388	2	089LH2
26	549.5	22.5	430	2	09A1I8
27	548.5	22.5	448	2	07N1U6
28	548.5	22.5	428	2	06DHF7
29	548	22.4	428	2	06WRT2
30	545.5	22.3	442	2	051880
31	542.5	22.2	666	2	07N1U3

32	540	22.1	655	2	07NKA9	07nka9 gloeobacter
33	538.5	22.0	667	2	08YT47	08yt47 anabaena sp
34	536.5	22.0	663	2	055107	055107 synchococc
35	534.5	21.9	667	2	055462	055462 synchocyst
36	531	21.7	404	2	088141	088141 pseudomonas
37	525.5	21.5	462	2	06N722	06n722 rhodopseudo
38	525	21.5	403	2	0883P3	0883p3 pseudomonas
39	525	21.5	446	1	NRTA SYN3	N73452 synchocyst
40	521.5	21.3	533	2	06W1F8	06w1f8 synchococc
41	510.5	20.9	459	2	0891A5	0891a5 bradyrhizob
42	509.5	20.9	559	2	07U3E5	07u3e5 synchococc
43	506	20.7	386	2	06N3I6	06n3i6 rhodopseudo
44	502.5	20.6	458	2	08YT49	08yt49 anabaena sp
45	494	20.2	402	2	0912V6	0912v6 pseudomonas

ALIGNMENTS

Query Match	Score	DB 2	Length
Best Local Similarity	65.4%	Pred. No. 6.3e-112	
Matches 301	Conservative 60	Mismatches 92	Indels 6
Gaps 4			
7 SSSKLLTSLASLAVMGLTIAPDVAGVLEKEDKPFKLTDMAPLAAAEKGFED 66			
6 ASPTLLLSAAATLA---LMLSSAQAAPLDVEDKQKGFILKTDMAPLAAAEKGFED 62			
67 EGIIFYOLEAQAQWVMDRVNNGELDGSMLAPAPLAASVGFQTKADIEVPSMGNGNA 126			
63 EGIYSTVLEPQAQWVKVLDLDRVIGELDGAHMLAGQFATIGFQTAQNVVTAASMDNGNG 122			
127 IYTSNEIYHOMKNPIL-BGKRVPHIKADYIKPVVEKTKAKGKPPNMAKPPASGSHNK 185			
123 IYTSNEVWERKPNLPKPGDKPLPHIKADALKPVAQYRAEGKPPKPMKVPFVSTHNYE 182			
186 LRYWLAAGINPGYVPPDIDSGQIGADALSVTPPPQMPSTLEAGTIFGYCVGEPMNOQ 245			
183 LRYWLAAGINPGYV-PDVDSGQIQADALSVTPPPQMPATLEAGTIFGYSGVEPMNOQ 241			
246 AVFKGIGVITDEELMKDTPKRVGVTYKQMAEKYPNTYLAATKALIRAAIMLADNNKN 305			
242 AVFKGIGVITDEELMKDTPKRVGVTYKQMAEKYPNTYLAATKALIRAAIMLADNNKN 301			
306 KREATIMLAQKQYGVADVLAASMGTEFEYKEDKRALPDPTNTPRHAASVPSYSAW 365			
302 RLEAAVYIKLSEVGVADAVIANSMGTREYKGDRAVDPDVFPRVYATVPYSDAW 361			
366 YLTOLRRQMINEFKPDNNYLDTAQNVYPPDIYLAAKELVVEGKAKADFP-ADTSIKP 424			
362 YLTQMRKQQLAEAKPDADYDETARKVYKPEIYLAARLLVEGKAKADFPMTSGCYFP 421			

QY 425 SONEFIDKVPEDANKPNDYLAKFAIGLKGKOTVAGKVV 463
DB 422 LDNGFIDGLAYDGRKRNVEYLTLPILGKGGAAVGGGLV 460

RESULT 2

ID Q7MDT5 PRELIMINARY; PRT; 475 AA.
AC Q7MDT5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nitrate transporter system, periplasmic component.
GN OrderedlocusNames=VVA0951;
OS *Vibrio vulnificus* (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of *Vibrio vulnificus*, a marine
pathogen."
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AF005348; BAC96977.1; -.
KW Complete proteome.
SQ SEQUENCE 475 AA; 53801 MW; B61BED1915DC367 CRC64;

Query Match 64.9%; Score 1586; DB 2; Length 475;
Best Local Similarity 65.6%; Pred. No. 4.4e-111;

Matches 292; Conservative 60; Mismatches 85; Indels 8; Gaps 5;

QY 20 LAWGLTTPADV--GAVGKLEKEDLKFGFIKLTDMAPLVAAEKGFEEBGLFVQLEAQA 77
DB 32 LTSW-LAFAPPVLAIEGEAEKEDLKFGFIKLTDMAPLVAAYKGFEEBGLVYTLLEAQA 90
QY 78 NMKVMDRVNGBLDSHMLAPAPLAASVGFQKADIEVPFSMGFNQNAITVSNEIWHQM 137
DB 91 NMVLLDRVIDGELDGAHMLAGQPLGATIGTQAVITAFSMDLGNNAITVSNVDWQOM 150
QY 138 KPNIPLEG-GKPVHPKADYLKPVVEKYKAEKGFPMNMTFPGSHNICKRWLAAGGIN 196
DB 151 KPMIAKOSDGKPVHPKADALKEVVTSYRDOGAFFMGWFPVSTHNYELRWLAAGGIN 210
QY 197 PGYSPPO-DISQIGADALLSVTPPOMPSTLEAGTIFGVCGEPMNOQAVFKGIGVPI 255
DB 211 PGYYAHKGDNSQINADVLSVTPPOMPATMEAGTIKGYCVGEPMNOQAVFKGIGVPI 270
QY 256 ITDEELMKOTPEKVFQVTKQMAEKYENTYLAATKALIRAAIWDADNNKRNKEAIEMLAQ 315
DB 271 VTDYEIMKNNPEKVFQVADQMAEKYENTHIRVVKALIRAHWLDENDNRRRAVAVKLSRS 330
QY 316 KOYVADVEVTLAASNMGTPEYEKDQKRALPDENTFFRHGASYSYSAYWYLTQLRRWGM 375
DB 331 SEYVGADAEVIANSMGTPEYEKDQKQVDFNVFFRYNATYYSYSDAIWYLTQMRWGM 390
QY 376 INEFKPDNMYLDTAKNVPYRPDIYLAAXELVAEGKAAEDFP--ADTSIKPSQNFIDK 432
DB 391 IEHQKDSWYMDLAKQVYRPDIYQRAAEALIEGTLISASDFPFAKESGRPRPQTHFIDQ 450
QY 433 VPPDANKPNDYLAKFAIGLKGKQTV 457
DB 451 IHYDGRSPNAYLQQFAIGLKGSESL 475

RESULT 3
Q8D6X8 PRELIMINARY; PRT; 487 AA.
ID Q8D6X8
AC Q8D6X8;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE ABC-type nitrate/sulfonate/dicarbonate transport system, periplasmic
component.
GN OrderedlocusNames=VZ0393;
OS *Vibrio vulnificus*.
OC Bacteria; Proteobacteria;
OC Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of *Vibrio vulnificus* CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016809; AA007349.1; -.
KW Complete proteome.
SQ SEQUENCE 487 AA; 54515 MW; FE5D5F1C5B865F4 CRC64;

Query Match 64.3%; Score 1572; DB 2; Length 487;
Best Local Similarity 65.1%; Pred. No. 5.2e-110;

Matches 289; Conservative 63; Mismatches 86; Indels 6; Gaps 4;

QY 20 LAWGLTTPADVGA-VGKLEKEDLKFGFIKLTDMAPLVAAEKGFEEBGLFVQLEAQA 78
DB 44 LTSWLAFAFPVLAIEGEAEKEDLKFGFIKLTDMAPLVAAYKGFEEBGLVYTLLEAQA 103
QY 79 NMKVMDRVNGBLDSHMLAPAPLAASVGFQKADIEVPFSMGFNQNAITVSNEIWHQM 138
DB 104 NMVLLDRVIDGELDGAHMLAGQPLGATIGTQAVITAFSMDLGNNAITVSNVDWQOM 163
QY 139 PNIPLEG-GKPVHPKADYLKPVVEKYKAEKGFPMNMTFPGSHNICKRWLAAGGINP 197
DB 164 PNIVDSOGKPVHPKADALKEVVTSYRDOGAFFMGWFPVSTHNYELRWLAAGGINP 223
QY 198 GYSPPO-DISQIGADALLSVTPPOMPSTLEAGTIFGVCGEPMNOQAVFKGIGVPI 256
DB 224 GYFAHKGDNQINADVLSVTPPOMPATMEAGTIKGYCVGEPMNOQAVFKGIGVPI 283
QY 257 ITDEELMKOTPEKVFQVTKQMAEKYENTYLAATKALIRAAIWDADNNKRNKEAIEMLAQ 316
DB 284 TDEYIMKNNPEKVFQVADQMAEKYENTHIRVVKALIRAHWLDENDNRRRAVAVKLSRS 343
QY 317 QYVADVEVTLAASNMGTPEYEKDQKRALPDENTFFRHGASYSYSAYWYLTQLRRWGM 376
DB 344 EYVGADAEVIANSMGTPEYEKDQKQVDFNVFFRYNATYYSYSDAIWYLTQMRWGM 403
QY 377 NEFKPDNMYLDTAKNVPYRPDIYLAAXELVAEGKAAEDFP--ADTSIKPSQNFIDK 433
DB 404 EHQKDSWYMDLAKQVYRPDIYQRAAEALIEGTLISASDFPFAKESGRPRPQTHFIDQ 463
QY 434 PEDANKPNDYLAKFAIGLKGKQTV 457
DB 464 RYDGRSPNAYLQQFAIGLKGSESL 487

RESULT 4

Q87HA7 PRELIMINARY; PRT; 454 AA.
ID Q87HA7
AC Q87HA7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative nitrate transport protein.
GN OrderedlocusNames=VPA1058;
OS *Vibrio parahaemolyticus*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=RIMD 2210633 / Serotype O3:H6;
 RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism
 RT distinct from that of *V. cholerae*.";
 RL Lancet 361:743-749(2003).
 DR EMBL; AP005087; BAC62401.1; -;
 KW Complete proteome.
 SQ SEQUENCE 454 AA; 50650 MW; CC1741FBA1E2D751 CRC64;
 Query Match 64.0%; Score 1564; DB 2; Length 454;
 Best Local Similarity 65.0%; Pred. No. 1.9e-109; Indels 8; Gaps 5;
 Matches 293; Conservative 62; Mismatches 88;
 QY 13 L L T L A S - S L A W M G L T I A P D V G A V G K L E K E D L K F I K L T M A P L A V A A E K F P E D S G L F V 71
 DB 6 M L K V S A I S L S V F - - T A V S A K A Q L G E P E I E D L K F I K L T M A P L A V A V E K F F E D S G L V 63
 QY 72 Q L E A Q A N K V M D R V V N G E L D S H M L A P A P L A S V G F G Y K A D I E V P S M G F N G A I T V S N 131
 DB 64 T L E A Q A N K V L I D R V I D G E L D G A H M L A G Q P L G A T I G V T K A E V I T A F S M D I N G A I T V S N 123
 QY 132 E I W H O K M P N I P L E - G G K P V P I K A D Y L K P V E K Y K A E G R P N M A M T P P A G S H N I K L R Y M L 190
 DB 124 D W E Q K P H I P K P D G P V P H I R A D S L V K P V D S Y R D K G S F N N G M V F P V S T H N Y E I R Y M L 183
 QY 191 A A G I N P G Y S P P Q - D I S Q I G A D A L L S Y T P P P M S T L E A G T I F G Y C V G E P M N Q A V F K 249
 DB 184 A A G I H P G F Y A P P S G N S G L D A D V L S T Y P P P M P A T M A G T I K G Y C V G E P M N Q A V F K 243
 QY 250 G I G V P I T D E E L M K D T P E K Y F G V T K O A E K Y P N T Y L A V T K A L I R A A I W L D A D N N K R K E A 309
 DB 244 G I G T P V T D E I M K N N B E K F G V S K A M E K N P M T H I R V Y K A L I R A A H M L D E N S A N R Q E A 303
 QY 310 I E M L A K O Y G A V E Y L A A S M N G T F E Y E K D K R A L D E N T F F R H G A S Y P S S A V Y L T Q 369
 DB 304 V K M L A S E Y V G A D A D V I A N S M T G T F E Y E K D K R D V P E N F F R H N T T P Y S D A I W Y L T Q 363
 QY 370 L R R M G M I N E F P D N W Y L D T A K N Y R P D I Y L A A K E L V A E G K A K A E D P P - - A D T S I K P S Q 426
 DB 364 M R M G Q I S S E K S D W M T D A V A K E Y R P D I Y Q A A Q A L I E D G V L S K D P P D S A D G R P P Q 423
 QY 427 N F F I D K V P P D A N K P N D Y L A K F A I G L K G Q T V 457
 DB 424 T H F I D N I V Y D G R E P N K Y L E K F S I G L K G K D K V 454
 RESULT 5
 QYUYV1 PRELIMINARY; PRT; 553 AA.
 AC QYUYV1
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Nitrate transport ATP-binding protein.
 GN Name=nrtc; Ordered locusNames=R8370;
 OS Rhodopirella baltica.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OC NCBI_TaxId=117;
 RN NCBI_SEQUENCE FROM N.A.
 RP STRAIN=1;
 RX MEDLINE=2275913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Bozsym K., Heilmann K., Rabus R.,
 RA Schlesner H., Aumann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete *Pirellula* sp.
 RT strain 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 DR EMBL; BX294133; CAD71540.1; -;

DR GO:0005524; F:ATP binding; IEA.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 553 AA; 60761 MW; 743DFAC3267B5D63 CRC64;
 Query Match 51.5%; Score 1257.5; DB 2; Length 553;
 Best Local Similarity 50.7%; Pred. No. 3.2e-86;
 Matches 249; Conservative 74; Mismatches 133; Indels 35; Gaps 8;
 QY 2 K T I I R S S K K L L T L T S A L A W - - - - - G L T I A P D V G A V G K - - - - - 36
 DB 62 K K A I R T N K L M L L G C A L A A V F T G C A D S G V S L E D P A A A K A V D I S K I E V D N P T S A A M L 121
 QY 37 - L E K E L K G F I K L T M A P L A V A A E K F F E D S L F Q L A Q A N K V M D R V V N G E L D S H 95
 DB 122 D L E K S L T G F I K L T D C A P L V I A K E G Y P D D E L A N T L E T Q S W K L L D N V I N G Q L D G A H 181
 QY 96 M L A P A L A S V G F T A D I E V P S M G F N G A I T V S N E I W H O K M P N I P - L E G K P V P I R A 154
 DB 182 M L A G P I G A T I G V T G S P I V T A Y S L D Y N G I T V S N E V A Q M Q E N D P E L K S P T P K P I S A 241
 QY 155 D Y L K P V E K Y K A E - G K P N M A M T P P A G S H N I K L R Y M L A A G I N P G Y S P P O D I S Q I G A D 213
 DB 242 A S L K P I V E D Y L Q D A G P P F M G M V F P V S T H N Y E I R Y M L A A S G I H P G M Y T - E S D I K G F T D A Q 300
 QY 214 A L L S T Y P P P M S T L E A G T I F G Y C V G E P M N Q A V F G I G V P I T D E L M K D T P E K Y F G T 273
 DB 301 V K L S T Y P P P M Q N L E A D I V K G Y C V G E P M N Q A V V T G I V P V T N D I M K N B E K Y F G Y T 360
 QY 274 K O M A E K Y P N T Y L A V T K A L I R A A I W L D A - - D N N K - - N R K E A I E M L A K O Y G A D V E Y L A A S 329
 DB 361 G E M A E K N P O T H A V I K A L I R A G K M L D A T D S G L V N R E A V E I L S K D Y V G A K E Y I G N S 420
 QY 330 M N G T F E Y E K D K R A L D E N T F F R H G A S Y P S S A V Y L T Q L R R M G M I N E F P D N W Y L D T A 389
 DB 421 M N G T F F O G S T D V R E M P D F V F F R H G A S Y P H S D A I F L T Q M R W G Q I T E S K P A S W A E T A 480
 QY 390 K A V Y R P D I Y L A A K E L V A E G K A K A E D P P A D - - T S I P S Q N F P I D K P P A N K P N D Y L A P 447
 DB 481 K K I Y K E I R O A A E L I S E G K L D P M E I P A P D Y G R A V A T T E P I D N K Y D A K P I G Y I N S F 540
 QY 448 A I G L K G Q T V A 458
 DB 541 E I G N K D D E S L A 551
 RESULT 6
 NRTC_SYNY3
 AC NRTC_SYNY3 STANDARD; PRT; 670 AA.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Nitrate transport ATP-binding protein nrtc.
 GN Name=nrtc; Ordered locusNames=sl11452;
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 OC NCBI_TaxId=1148;
 RN NCBI_SEQUENCE FROM N.A.
 RP STRAIN=1;
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasegawa S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nartuo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -I- FUNCTION: Probably part of a high-affinity binding-protein-
 CC dependent transport system for nitrate. Probably responsible for
 CC energy coupling to the transport system.
 CC -I- SUBCELLULAR LOCATION: Membrane-associated (Potential).

DT 05-JUL-2004 (Tremblrel. 27, last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, last annotation update)
DE AGR_L1886P.
GN OrderedLocusNames=AGR_L1886;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Eubacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorlo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Tarchonk O., Epp A., Liu F.,
RA Flanagan C., Allinger M., Doughty D., Scott C., Lampas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; A5008293; AAK89514.1; -
DR InterPro; IPR009073; HSC20_C.
DR InterPro; IPR006311; Tac.
DR TIGRFBMS; TIGR01409; TAT_signal_seq.1
SQ SEQUENCE 471 AA; 51726 MW; 7650A568C39D79A CRC64;
.
Query Match 25.7%; Score 627.5; DB 2; Length 471;
Best Local Similarity 35.5%; Pred. No. 8.7e-39;
Matches 166; Conservative 70; Mismatches 171; Indels 61; Gaps 19;
QY 1 MKTITSS-SKULL--TISASLAWGLTAPD--VGAVKLEKEDLKGFILDTMAPL 55
DB 42 MKKIFSGTYSRRLITLTTAALVTAVRTAFPSGAPATAAPEVKAKGFIATLDAAPL 101
QY 56 AVAAEGFPEDEGL-FVQLEAOANMKVMDRVVNG---ELDSHMLAPAPLAASVGF 110
DB 102 IIAAEGLFPAKGMPEVEYLKQSKGATRDNLVLGASNGIDCAHILTPPYILMTGKYT 161
QY 111 KADIEVPS---MGFNNAITVSNIEIHWQMKNPIDLEGKPVHPKADYLKVEVEKYA 166
DB 162 QNNVPVPMITLARNLNDSGQISVAK---YAEFGVGLDASK-----LKAAPFKKKA 209
QY 167 EGKPFMAATFPAGSNIKIRYLAAGINPGYSPDIDSGIDGADLISVTPPOMPS 226
DB 210 DGEIIRKAAATFPFGTDLIRYLAAGID---PDKQVS-----TIIVPPQOMA 256
QY 227 TLEAGTIFGCVGEPNNOQAVPFKIGVPIITBELAKDTPKGVGTAKMAEKYPTTYLA 286
DB 257 NMKVGMDVFCVGEPMNEQLVNOGIFCTACTGELMKHPEKALGRADOMEKNPATYA 316
QY 287 VTKALIRAAIWLADANNKREKAEIEMIAOKYGVADVEVLAASNGTFEYKDKRALPD 346
DB 317 LMAVVEAQOM--CDSEMANKEEMSTILGRQMFNPVPKQVGLKKNINY--GNGVLE- 371
QY 347 FNT-----FRRGASIPYSISAVVYLTQLRRGMINEFKEDMNYLDTAKNVRPDIYLA 401
DB 372 -NTGLQMKFWODIASYPFHSHDSWFTENIRWG--KFAPDITVAKLVAKVNEDITWRAA 427
QY 402 AKELVAEGAKADEPDAITSIKPSQNFIDKVPFANKNDYLAARAI 449
DB 428 AKDL--GVA--DLPASTS-RKKEFTFGKV-FDEPNPSAYLESLSI 467

OS Rhizobium meliloti (sinorhizobium meliloti).
OG Plasmid pSymB.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhölter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiotic Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL591985; CAC49093.1; -
DR PIR; E95928; E95928.
KW Complete proteome; Plasmid.
SQ SEQUENCE 441 AA; 47208 MW; 2F2F4498B8F81D58 CRC64;
.
Query Match 25.2%; Score 616.5; DB 2; Length 441;
Best Local Similarity 33.8%; Pred. No. 5.4e-38;
Matches 153; Conservative 74; Mismatches 182; Indels 43; Gaps 12;
QY 19 SLAWGLTAPDVAVGKLEKEDLKGFILDTMAPLAANAKGFPEDEGLVQLEAOAN 78
DB 19 NLSAGALPAP--GSIRHSGSRTRAGFIPLDAVTLAAAEFGAORRGITLTVKQVS 76
QY 79 MKVMDRVVNGELDSHMLAPAPLAASVGFKTADIEV--PFSMGFNNAITVSNIEIHWQ 137
DB 77 WANVRDLAFROPDVAAHLSMPVVAAMLGSGNSPSTIPFSLGSGNAITSTRLYGLM 136
QY 138 KNPILEGKPVHPKADYLKVEVEKYAEGK--PNNMANTFPAGSNIKIRYLAAGIN 196
DB 137 OQAGGIGGEDALK--NAKALAIVIRSAGAAGRPPLTIGVTFPSSHNYEFRTYLAAGID 195
QY 197 PGYSPDIDSGIDGADLISVTPPOMPSSTIEAGTIFGCVGEPNNOQAVKIGVPI 256
DB 196 P-----DRDVLVVPPMPTSDALAAIGIDFCVGAPEWNAVASSRGVRIYA 242
QY 257 TDEELKDTPEKVPFGVTKMAEKYPTTYLAATVTKALIRAAIWLADANNKREKAEIEMIAOK 316
DB 243 TKQDITPSPAPPEVITGMREPMAEANDTVSRVLVALLDRARW--SDEPANKGLAEVLAEE 300
QY 317 QYGVADVEVLAASNGTFEYKED--DKRALPDENTFRRGASYSYSAAVVYLTQLRRWG 375
DB 301 RHVAAPADIIIRVLAGEFAIDPEGNRVVENYLVFHAGFANYPRSQALMTYSQWVRWG 360
QY 376 INEFKPDNMYLDTAKNVRPDIYLAAKEVLAEGAKADEPDAITS-----KPSQNFIT 430
DB 361 TSLSRQ--RVDAAISAYRPDLRYREAL-----GK---DALPADADAGRLBGATDGRPM 408
QY 431 DKVPFANKPNNDYLAKFAIGLKQTVAGGV 462
DB 409 DGHVDFPGRIEDYIKSF-----GAPSVNRSV 435

RESULT 9
Q92VK8
ID Q92VK8 PRELIMINARY; PRT; 441 AA.
AC Q92VK8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, last annotation update)
DE Putative nitrate transport protein.
GN ORFNames=SMB21114;

RESULT 10
Q8U924
ID Q8U924 PRELIMINARY; PRT; 431 AA.
AC Q8U924; Q7CT12;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, last annotation update)
DE ABC transporter, nucleotide binding/ATPase protein
DE (AGR_L1881g1p).
GN Name=ntc; OrderedLocusNames=AGR_L1881g1, Actu3906;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Serubal J.C., Kaul R., Monte D.E., Kitzajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D.S.,
RA Chapman P., Clendenning J., Decherage G., Gilllet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-X., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RT Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Currolo B., Goldman B.S., Cao Y., Akenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iarchouk O., Egg A., Liu P.,
RA Wollem C., Allinger M., Dougan D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gureon J., Lomo C., Seer C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RT Science 294:2323-2328(2001).
RL EMBL; AE009321; AA144714.1; -
DR EMBL; AE008293; AA89513.1; -
DR PIR; AD3037; AD3037.
DR PIR; G98248; G98248.
KW Complete proteome.
SQ SEQUENCE 431 AA; 46254 MW; 22DB764D66992C63 CRC64;

Query Match 25.0%; Score 611.5; DB 2; Length 431;
Best Local Similarity 35.5%; Pred. No. 1.2e-37;
Matches 151; Conservative 69; Mismatches 166; Indels 39; Gaps 12;

QY 34 VGLKEKEDLKFGRFKITDPAFLVAAEKFEDEGLFVLEAQAANKVMDRVNSELNG 93
DB 24 VSGDRKRIKAGFTPLVDASVLIAAAGFGADREGILDLVDKVSANVRDLAERQFDI 83
QY 94 SHMLAPAPLAASVFGTKADIEV-PSMGEFGNAITVSEIWMKPNIP-L-EGKRPVAP 151
DB 84 AHWLSMPVASMGLGNSPSTITPFSLRGSAIITLSTRKFRMALKGLSTAGALEN 143
QY 152 IKADYIKPVVEKYKABGK-PENNAWTFPAGSHNIKLRWYLAAGINPGYSPQDISQI 210
DB 144 ARA--LKLVLDDMRARAGEAPPTLGMTYPFSSHNYFRWYLAAGIHPDH----- 190
QY 211 GADALLSVPPPMPTLEAGTFCGYCVGPMNQAVFVGIGVPTVDEIMKDPREKVF 270
DB 191 --DVKLVVPPPTSDALAGALDGCVCVGPVNIVAERGVGIVAKKDLPMSAEKVI 248
QY 271 GYTKQWAEKYPNTYLAFTVALLRAATWLD-ADNNKRRKEAIEMLAQKQYVGADEVITAA 329
DB 249 GMRPEWABEQEIVGVLTLALDAASWCDLADHDLDSG--LADPRITGAQPSIIRV 305
QY 330 MNGTFEYE-KDDKRALPDENTFPRHGAASPSYSAAVWYLTQLRRWGMINEFKDWMYLD 388
DB 306 LAGEFISIDSGNRRVIEKFTFPHGDANYPROSQSLIMYSQIMRWG--QAELSETGVNA 362
QY 389 AKNVVPPDIYLAAKELVAEGKAKADPADPSIK-----PSQNFIDKXFPDANKNDL 444
DB 363 ALSAVPDIYRAA-----LGDGKA-----PEDADIRIEGDEGRFVDGVPDPADTAGV 413
QY 445 AKFAI 449
DB 414 NSPAV 418

```

RESULT 11

```

NRPA ANASP
ID NRPA ANASP STANDARD; PRT; 440 AA.
AC 044262; 006469;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nitrate transport protein nrta.
GN Name=nrta; OrderedLocustNames=nrta;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97144534; PubMed=8990301;
RA Frias J.E., Flores E., Herrero A.;
RT "Nitrate assimilation gene cluster from the heterocyst-forming
RT cyanobacterium Anabaena sp. strain PCC 7120."
RT J. Bacteriol. 179:477-486(1997).
RN [2]
RP REVISIONS.
RA Frias J.E.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matano A., Iriuguchi M., Ishikawa A., Kawashina K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
RN [4]
RP SEQUENCE OF 1-92 FROM N.A.
RX MEDLINE=97136629; PubMed=8982006;
RA Cai Y., Wolk C.P.;
RT "Nitrogen deprivation of Anabaena sp. strain PCC 7120 elicits rapid
RT activation of a gene cluster that is essential for uptake and
RT utilization of nitrate."
RL J. Bacteriol. 179:258-266(1997).
CC -I- FUNCTION: Essential component of the nitrate-transporting system.
CC -I- SUBCELLULAR LOCATION: Inner membrane-associated (potential).
CC -I- SIMILARITY: Strong, to carotenoid-binding protein A (cbpA).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X99709; CAB8041.2; -
DR EMBL; AP003583; BAB72566.1; -
DR EMBL; U61496; AAC46075.1; -
DR PIR; AG1882; AG1882.
DR InterPro; IPR006311; Tac.
DR TIGRfam; TIGR01728; Seta_fam.1.
DR TIGRfam; TIGR01409; Tat_signal_seq.1.
KW Complete proteome; Inner membrane; Nitrate assimilation; Transport.
FT CONFLICT 100 W -> C (in Ref. 1).
SQ SEQUENCE 440 AA; 48475 MW; 29937A41FB45CE9C CRC64;

Query Match 25.0%; Score 610.5; DB 1; Length 440;
Best Local Similarity 33.4%; Pred. No. 1.5e-37;
Matches 158; Conservative 71; Mismatches 173; Indels 71; Gaps 17;

QY 8 SSKLLLT-----LSASLAWMGITL-----APDVGANG--LKEKEDLKF 44
DB 5 SRRKPLFTTGAAMASILVHGCTNSGOSATTEQAPSAAPAVANVSANAPRVETTKAKL 64

```

QY	45	GIKLTLDMAPIVAAEKEGFEDEGE.F - VOLEQA0AMKVMDEV----	-VNGELDGSNHLAP	99
Db	65	GRITPLTDAAPLIIAEKEGFFAKYGMIDIEVYQKSPVYTRDLTKIGSSGGGIDGAHILSP		124
QY	100	APLAASVGEFTADIIEVPFSMGFNALITSNEIWHOMKENIPILEGKRVHPIKADYLKP		159
Db	125	MPYLMTI--NDKVPYVILIARLMTNQGALISAAKFP---KELVNLE-----SKSLKD		170
QY	160	VEKTKYAEKGPPNNAMTTPPAGSHNITKLRYWLAAGINPEGYSPPODISGQIGADALLSVT		219
Db	171	LAIKKAKDKALKMGITTFPGCTHDLMRWYMLAAGINP-----DQDVLEAV		217
QY	220	PEPQMPSTLEACTIGYCVGSEPMNOQAVPKGIVGVITDELMKDPTEKVPSTQWMAEK		279
Db	218	PEPQVAVNNKNTVGVFCVGEPPMAQVLVQKIGYSALVTGELMKHPKPAFASMRQDWLEQ		277
QY	280	YENTLAVTKALIRAIWLDADNNKKRKEAIEMLAQOKYGVADVEVLAASNNGTREYEKD		339
Db	278	NPNAQALIMALLEAQOM--CDKANKKEMCKICSDRKYNVAADIIERAGNIDY--G		333
QY	340	DKRALPDF--NTFPRHGASYSYSAAVWYITQLRRMGMINFEKPDNYLDTAKNVYPD		396
Db	334	DGRKQNFNAHRMKFWADNASYPYKSHDIFWLTEDIREDIRMGVLPK---DTKQDIVNOVKED		390
QY	397	ITYLAAKELVAEGKAKAEDPFDSTISQSNFIDKVPDPANKPDYLAKEFI		449
Db	391	LMKKAK--AIGVADAE-IPASSS-RGVETFF-DGVKFPDKPEEYINSLKI		437
RESULT 12				
Q7NEC4				
ID	Q7NEC4	PRELIMINARY;	PRT;	429 AA.
AC	Q7NEC4;			
DT	01-MAR-2004	(TREMBLrel. 26, Created)		
DT	01-MAR-2004	(TREMBLrel. 26, Last sequence update)		
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)		
DE	G113956	protein.		
GN	Ordered locus names=g113956;			
OS	Gloeobacter violaceus.			
OC	Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.			
OX	NCBI_TaxID=33072;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=PCC 7421;			
RX	MEDLINE=22977040; PubMed=14621292;			
RA	Nakamura Y., Kaneko T., Sato S., Miumuro M., Miyashita H., Tsuchiya T.,			
RA	Saenoco S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,			
RA	Kohara M., Matsunoto M., Matsuno A., Nakazaki N., Shimo S.,			
RA	Takeuchi C., Yamada M., Tabata S.;			
RT	"Complete genome structure of Gloeobacter violaceus PCC 7421, a			
RT	Cyanobacterium that lacks thylakoids";			
RL	DNA Res. 10:137-145(2003).			
DR	EMBL: AP006581; BAC91897.1; -.			
KX	Complete proteome.			
SO	SEQUENCE 429 AA; 47153 MW; 205101AE5C8B22F9 CRC64;			
Query Match 25.0%; Score 610; DB 2; Length 429;				
Best Local Similarity 35.4%; Pred. No.1,6e-37;				
Matches 162; Conservative 60; Mismatches 174; Indels 62; Gaps 14.				
QY	16	LSASIAVWGLTI-ADVGAVGKLEKEDIKFGIKLTLDMAPIVAAEKGFDEGLFVLE		74
Db	12	IAGAAATIGPTLMLPKALAKGLEPKLVGFIALIDCAPVIAERGGFERHGDVLELS		71
QY	75	AQAMNKVMDRVVNELDGSNHLAARPLAASVGFETK--ADIEVFPMSGFNALITVNE		132
Db	72	KEESVAVSREGILTRLDASHALAGTPIAVQVAGABKAPARLITAMSLDINGNATTFSKR		131
QY	133	IWHQKKPNIPILEGKRVHPIKADYLKPVVEKYKAEKGPPNNAMTTPPAGSHNITKLRYWLA		192
Db	132	LW-----QAGVRSG-----AD-LKIIATIGKV-GRITLGMAMVASSMYNNLCYWLAH		177
QY	193	GGINFGYVSPDIDSGQIGADALLSVTPPQMPSTLEAGTIFGVCVGEPPMAQAVFKIG		252

```

Db      178 HDLHP-----YRDVRLTLTPRQGLIALEBGNIDFCVTEPPNTRHAGAG 224
QY      253 VPVITBELMKTPREYFGVTKQAKAEKPNITYLATKTLTAALWLDADNNKRAEIAM 312
Db      225 FTVALARDWGGHPEPEVLAVMPEPMARTEPHNLALVKLLACASY--CDEPAHQEVLRT 282
QY      313 LAQKQYVGADVETLAASNNGTFEYEKDD-----KRALPDENTFPRHAGASY----- 357
Db      283 LSPRTYLAQRFLEBPALESLGRDLDFGFGQEGEKATSLTDKVDPEFVFFPRKDDYLVGNDQA 342
QY      358 --SSYSAAWVYLTLQRLRMGMINPEKPD--NWYLDPAKQNYRRPDYLLAAKEELVAGSKAAE 414
Db      343 TPFMKSHGMLTLTQMARNGQIPAIPIAGVDKLLD---RYRRDITYRQAAAEIGT--KAPAQ 397
QY      415 DFPADTSIKPSQNFIDKVPDPANKPNDYLLAKFAIGLK 452
Db      398 DY-----KSENTFTDRRRFDPSPNPVAYLVSFEELIAR 428

```

ID	Q7NKB1	PRELIMINARY	PRT:	427 AA.
AC	Q7NKB1			
DT	01-MAR-2004 (TREMBlrel. 26, Created)			
DT	01-MAR-2004 (TREMBlrel. 26, Last sequence update)			
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)			
DE	Nitrate/nitrite transport system substrate-binding protein.			
GN	Name=nitA, Order=locusNames=91r1567;			
OS	Gloebacter violaceus.			
OC	Bacteria; Cyanobacteria; Chroococcales; Gloebacter.			
OX	NCBI_TaxID=33072;			
RA	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=PCC 7421;			
RX	MEDLINE=22977040; PubMed=14621292;			
RA	Nakamura Y., Kaneo T., Sato S., Mimuro M., Miyashita H., Teuchiya T.,			
RA	Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,			
RA	Kohara M., Matsumoto M., Matsuo A., Nakazaki N., Shimo S.,			
RA	Takeuchi C., Yamada M., Tabata S.,			
RT	"Complete genome structure of Gloebacter violaceus PCC 7421, a			
RL	cyanobacterium that lacks thylakoids.";			
DR	DNA Res. 10:137-145(2003).			
DR	EMBL; AP006573; BAC89508.1; -.			
DR	InterPro; IPR006311; Tat.			
DR	TIGRPFMS; TIGR01409; Tat_signal_seq; 1.			
DR	Complete proteome.			
SO	SEQUENCE 427 AA; 46542 MW; COD86F1625E39273 CRC64;			
QY	Query Match	24.3%;	Score 594;	DB 2; Length 427;
	Best Local Similarity	33.8%;	Pred. No. 2,66-36;	
	Matches 152; Conservative	73;	Mismatches 169;	Indels 56; Gaps 14;
QY	9 SKLLLTLSASLAWGLTIADVGAVGLKEKEDLKFGFIKLTMDAPLAAAEKGFEEDEG 68			
DB	11 TRRGILQAAAAATGSIITPGYGAADAPETTKARLGISLSDCAPLIAAEKGLFDRYG 70			
QY	69 L-FVQLAQAQNMKVMRNV---NGELDGSNMLAPPLAASVFGTKADIEVFP-----S 119			
DB	71 MKDEVAVKQASWGVTRDNLLEAGGGGIGDAHILTPMIVYLIIANGNITKSGSKVPEFILIAR 130			
QY	120 MGFFNGNATITVSNELTWHQMKPNIPLGGGKPVHPIKADYILKPVVEKYKAEKGFNNAMTFPPA 179			
DB	131 LNVNGGGSIVANK--YTL-----LKVGLDAAPMIAEALKA-----KANGDITTYAQTFFPG 178			
QY	180 GSHNIKIRYWLAAAGINPGYYSPPODISGOIGADALLSVTPPOMPSTLEAGTIFGYCVG 239			
DB	179 GTTHAMIRYWLAAAGIDP-----ETDVKMITVPPPPQVMNMKGTVDAFCVG 225			
QY	240 EPMNQQAQVFGIGPVITDEIMKDPTEKVPFGVTKQMAEKYPNTYLAVTKLIAAATMLD 299			
DB	226 EPMHQDLINQDGYTAVTGTQIMRHPKESFALRADYVERKPKATKALLMVAQEAQIW-- 283			

```
QY 300 ADNNKRAKEIEMLAOKOYVADVEVLAASMGTFEY--EKDDKRALPDFTFPRHGAAY 357
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 284 ADKAEKMDLAQYVRSWATGAPVSDI VARYKGIIDYGDGRPEYRSPHMQWKPABAY 343
QY 358 PVSYSAVWYLTQLRRMGMINFEFPDMWYLDYAF--NVYRPDIYLAAKELVAEGKAKAE 414
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 344 PYOSHLIMWLTEDIRWGVLP A-----TTDTKLVAAVNRNEDIMREAAKAL-----CQ 390
QY 415 DFPADTSIKPSQNFIDKVPEDANKPDY L 444
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 391 PAPKGTIS-RGVEKFF-DGVAFDPTKEBAY L 418

RESULT 14
Q8Y276 PRELIMINARY; PRT; 657 AA.
ID Q8Y276;
AC Q8Y276;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Nitrate transport ATP-binding protein.
GN Name=nrtc; OrderedlocusNames=aiT0610;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_Taxid=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneo T., Nakamura Y., Molk C.P., Kuritz T., Sasaemoto S.,
RA Watanabe A., Itiguchi M., Ishikawa A., Kawashima K., Kikura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RC DNA Res. 8:205-213(2001).
CL -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF003583; BAB72568.1; -.
DR PIR; A11882; A11882.
DR HSSP; OS6206; 1L2T.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0015112; F:nitrate transporter activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0015706; P:nitrate transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfam; TIGR01184; ntrCD; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 657 AA; 73294 MW; F7E1434798109324 CRC64;

Query Match 24.0%; Score 587; DB 2; Length 657;
Best Local Similarity 35.8%; Pred. No.1.5e-35;
Matches 150; Conservative 74; Mismatches 139; Indels 56; Gaps 17;

QY 39 KEDLKEGFIKLTDMAPLAAVAAEKGFPEDEGL-FVQLEAQAAMKRVMDRVNNGELDSHNL 97
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 276 KAVIEIGFMPLTDSAPLI VAKKEGFAKGLDNVILIRANNMQAINTGVVTKGLDAAGV 335
QY 98 AAPPLAASVGFQIKALIEV--PSSMGNGALITVSNIEIHQMKPNITPLESGKRVHPIKD 155
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 336 AGPMLALTIGAGSQGTPTPVINALNLSRNNAAITFSKRLYNQGVRS L-----AD 383
QY 156 YLKPVVEKKAEGKPFNMAMTTPAGSHNITKLRYWLLAAGINPGVSPPODISQIGADAL 215
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 384 -LKAVID--SSPQIITLIGVSHASMONLRLRYWLLAAGID-----PDSVDS----- 427
QY 216 LSVTPPPQMFSTLEAGTIFGYCVGEPPMNOAVFKGIGVPIITDEILMKDTPKRVGVTKQ 275
```

```
Db 428 LTVIPPTQVQSLQKAGNIDGYCAGEPPMWYQAAYHDDLGFVAATALFIMSQPKKVLGVARD 487
QY 276 WAEKTPNTYLAATKALLIRAITWLDADNNKRAKEIEMLAOKOYVADVEVLAASMGTF- 334
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 488 WAKYPEPTYLNVKALIEACKY--CDLIRNREIIEILCRPEY----LDVNAAYVRSGR 541
QY 335 -EYEKD--KRALDFTNTPFRHGAYSYSASVWYLTQLRRMGMINFEFPDMWYLDYAF 390
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 DRYRGGDGTTPPOLATVNOFYLNKTYNTPRTIIMITOMAWG-LTPF-PQW-VETIE 598
QY 391 NVYRPDIYLAAKELVAEGKAKAEDEPADTSIKPSQNFIDKVPEDANKPDY LAKFAI 449
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 599 RVCRTDIFGAARDL-----GLLDIGEDDP I---HLFDGKL-FNPSPIEYLSLEI 646

RESULT 15
Q8Y2F5 PRELIMINARY; PRT; 420 AA.
ID Q8Y2F5;
AC Q8Y2F5;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE PUTATIVE NITRATE TRANSPORTER PROTEIN.
GN Name=naef; Synonyms=RS03349; OrderedlocusNames=RS03381;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EM11000;
RC MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Broctier P., Camu J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,
RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646058; CAD13909.1; -.
KW Complete proteome.
SQ SEQUENCE 420 AA; 45516 MW; 6BC0BDBABF82F3C CRC64;

Query Match 24.0%; Score 585.5; DB 2; Length 420;
Best Local Similarity 32.8%; Pred. No.1.1e-35;
Matches 151; Conservative 82; Mismatches 164; Indels 63; Gaps 15;

QY 3 TIRSSSKLLTLLS--ASLAWGGLTIAPDVGAVGK--LEKEDLKEGFIKLTDMAPLAAV 58
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 TINKRRRRVLAIVAGSGAMALIDPLVYAGAAASDAPBEKELKVGFIPLTDCASVWA 68
QY 59 AEKGFPEDEGLFVQLEAQAAMKRVMDRVNNGELDSHMLAAPPLAASVGF-GTKADIEVP 117
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 ATLGDKKYGKIKIVSKEASWAGVADKLVSGLDAHYLVGLVGVQIGGPKDMAY L 128
QY 118 FPMGNGNALITVSNIEIHQMKPNITPLESGKRVHPIKADYLKRVVAKRYAEGKPFNMAMT 177
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 MTLNNGAATITLSSKL-----KEAGVDSGLKALMTR-----EKDYIYFAQTF 172
QY 178 PAGESHNIKRLRYWLLAAGINPGVSPPODISQIGADALLSVPPQMFSTLEAGTIFGYC 237
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 173 PGTGHAAMLTYWLLAHGHP-----LQDAKAITYPPFOVAMNRGVAMNDGYC 219
QY 238 VGEPPMNOAVFKGIGVPIITDEILMKDTPKRVGYTKQWAEKTPNTYLAATKALLIRAIT 297
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 220 VGEPPGARAIDAGIGFTLETTOAIWKDHPKVLGTTAFAQYKVPNTAVALTAIVLEASKF 279
QY 298 LDADNNKRAKEIEMLAOKOYVADVEVLAASMGTFEY-----EKDDKRALPDFTF 350
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 280 IDA-SASNRKRTAETVAASVYNTMDIILDRMLGRYTNGLGKTWDADPMR-----F 331
```


Qy 351 FRHGA-SYPESSAVWYLTOLRRWGMINEFKPDNMWYLDTAQNYYRPDIYLAAKEIYAEG 409
Db 332 YHDAVNFYLSIDGMWFLIQHKEWGL-KTHPD-YLAIAQVNRVDIYKQAA----- 381
Qy 410 KAKAEDFPADTSIKPSQNFIDKVPDPANKPNDYLAKFAI 449
Db 382 AATGTPLP-KSDLRJAR-LIDGVWMDAKNPAAYADSFKI 418

Search completed: September 16, 2005, 01:09:11
Job time : 181 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 00:38:09 ; Search time 44 Seconds

(without alignments)
787,208 Million cell updates/sec

Title: US-10-689-200-2

Sequence: 1 MKTIISSSSKKLLTISASU.....AKFAIGKKQIVAGKVD 464

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A COMB.pep:*
2: /cgn2_6/prodata/1/aa/6A COMB.pep:*
3: /cgn2_6/prodata/1/aa/6B COMB.pep:*
4: /cgn2_6/prodata/1/aa/6B COMB.pep:*
5: /cgn2_6/prodata/1/aa/6B COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	577.5	23.6	456	US-09-489-039A-10363	Sequence 10363, A
2	446.5	18.3	474	US-09-540-236-2892	Sequence 2892, Ap
3	338.5	13.9	336	US-09-328-352-7572	Sequence 7572, Ap
4	292.5	12.0	249	US-09-252-991A-30501	Sequence 30501, A
5	240	9.8	419	US-09-252-991A-25707	Sequence 25707, A
6	219.5	9.0	275	US-09-602-787A-384	Sequence 384, App
7	157	6.4	197	US-09-252-991A-30392	Sequence 30392, A
8	121.5	5.0	336	US-09-902-540-12264	Sequence 12264, A
9	116.5	4.8	623	US-09-252-991A-26757	Sequence 26757, A
10	116.5	4.8	623	US-09-252-991A-26757	Sequence 26757, A
11	103.5	4.2	327	US-09-489-039A-12136	Sequence 12136, A
12	102	4.2	369	US-09-107-532A-5754	Sequence 5754, Ap
13	102	4.2	1062	US-09-397-550-4	Sequence 4, App1
14	102	4.2	1076	US-09-470-443-6	Sequence 6, App1
15	102	4.2	1082	US-09-397-550-5	Sequence 5, App1
16	102	4.2	1109	US-09-397-550-6	Sequence 6, App1
17	102	4.2	1115	US-09-397-550-23	Sequence 23, App1
18	102	4.2	1145	US-09-470-443-2	Sequence 2, App1
19	102	4.2	1145	US-09-470-443-4	Sequence 4, App1
20	102	4.2	1145	US-09-397-550-20	Sequence 20, App1
21	101.5	4.2	952	US-09-328-352-5611	Sequence 5611, Ap
22	101.5	4.2	1252	US-10-012-762-2	Sequence 20, App1
23	101.5	4.2	1252	US-09-704-036B-20	Sequence 20, App1
24	101	4.1	858	US-09-255-829-22	Sequence 22, App1
25	101	4.1	858	US-09-255-829-29	Sequence 29, App1
26	101	4.1	1169	US-09-255-829-20	Sequence 20, App1
27	98.5	4.0	315	US-09-248-796A-16091	Sequence 16091, A

28	98.5	4.0	1864	2	US-08-804-227C-3	Sequence 3, App1
29	98	4.0	324	4	US-08-956-171E-5209	Sequence 5209, Ap
30	98	4.0	324	4	US-08-781-986A-5209	Sequence 5209, Ap
31	96.5	4.0	448	4	US-09-198-452A-216	Sequence 216, App
32	96.5	4.0	457	4	US-09-438-185A-199	Sequence 199, App
33	96.5	4.0	659	2	US-08-258-639A-4	Sequence 4, App1
34	96.5	4.0	659	2	US-08-900-951-4	Sequence 4, App1
35	96.5	4.0	659	2	PCT-US95-07391A-4	Sequence 4, App1
36	95.5	3.9	893	4	US-09-489-039A-14127	Sequence 14127, A
37	93.5	3.8	540	4	US-09-902-540-14003	Sequence 14003, A
38	93.5	3.8	1018	1	US-08-452-052-2	Sequence 2, App1
39	92	3.8	620	3	US-09-442-100-11	Sequence 11, App1
40	92	3.8	620	3	US-08-939-105-11	Sequence 11, App1
41	92	3.8	620	4	US-09-442-102-11	Sequence 11, App1
42	91	3.7	379	4	US-09-071-035-6	Sequence 6, App1
43	91	3.7	402	4	US-09-489-039A-11633	Sequence 11633, A
44	91	3.7	431	4	US-09-134-000C-6307	Sequence 6307, Ap
45	91	3.7	777	4	US-09-917-254-57	Sequence 57, App1

ALIGNMENTS

US-09-489-039A-10363	RESULT 1	US-09-489-039A-10363	Application US/09489039A
Sequence 10363	Patent No. 6610836	Sequence 10363	Application US/09489039A
GENERAL INFORMATION:	APPLICANT: Gary Breton et. al	GENERAL INFORMATION:	APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA	TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA	TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA	TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001	FILE REFERENCE: 2709.2004001	FILE REFERENCE: 2709.2004001	FILE REFERENCE: 2709.2004001
CURRENT FILING DATE: 2000-01-27	CURRENT FILING DATE: 2000-01-27	CURRENT FILING DATE: 2000-01-27	CURRENT FILING DATE: 2000-01-27
PRIOR FILING DATE: 1999-01-29	PRIOR FILING DATE: 1999-01-29	PRIOR FILING DATE: 1999-01-29	PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342	NUMBER OF SEQ ID NOS: 14342	NUMBER OF SEQ ID NOS: 14342	NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10363	SEQ ID NO 10363	SEQ ID NO 10363	SEQ ID NO 10363
LENGTH: 456	LENGTH: 456	LENGTH: 456	LENGTH: 456
TYPE: PRT	TYPE: PRT	TYPE: PRT	TYPE: PRT
ORGANISM: Klebsiella pneumoniae	ORGANISM: Klebsiella pneumoniae	ORGANISM: Klebsiella pneumoniae	ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10363	US-09-489-039A-10363	US-09-489-039A-10363	US-09-489-039A-10363
Query Match	Query Match	Query Match	Query Match
Best Local Similarity 23.6%; Score 577.5; DB 4; Length 456;	Best Local Similarity 23.6%; Score 577.5; DB 4; Length 456;	Best Local Similarity 23.6%; Score 577.5; DB 4; Length 456;	Best Local Similarity 23.6%; Score 577.5; DB 4; Length 456;
Matches 153; Conservative 76; Mismatches 173; Indels 55; Gaps 17;	Matches 153; Conservative 76; Mismatches 173; Indels 55; Gaps 17;	Matches 153; Conservative 76; Mismatches 173; Indels 55; Gaps 17;	Matches 153; Conservative 76; Mismatches 173; Indels 55; Gaps 17;
QY	7	SSSKKLLTISASIAVGLTIPDV-----GAVGKLEKEDLKFGFTKLTDMAPLVAAR	60
DB	44	SISRRLDGAALG--GAMLLPGVQAAAGSGDKPEQRYRVGFTPLTDCAPLAIANA	101
QY	61	KGFEEDEGLFVQLEAQAANKVMDRVVNGELDSHMLAPLAASVFGTKADIEVP-PS	119
DB	102	KGGDQKYGITIVAAKSAKMAAARDKLVAGELDAHLVGLVGLGELGASKPMANMT	161
QY	120	MENGNATIVSNEIWMQKPNIPBGGKRVHPKADYLPVVEKRYAEGKPEMNMTFPA	179
DB	162	LNNNGQATLISSEL--QEGKVTDLG-----LKRLLDR-SAPSS-YFATFPFL	206
QY	180	GSHNITLRLVWLAAGINPGYSPPODISGIGADALLSTPPROMSTLEAGTIFCYCG	239
DB	207	GTHAAMLVYWLASAGIDP-----FNDVTVVPPQVMNRIGMSGFCVG	253
QY	240	EPNQAQAVFKGIGVPVITDEBELKQTPKRVFGVTKQMAEKYPTVLAATKALIRAIWLD	299
DB	254	EPNNAALINDRIGFTAAISQDIWPEHEKVLGTRBRDVRNRTAALVAALMEAGRWI-	312
QY	300	ADNNKRAKAIEMLAQKQYVGADEVTLASNGTPEYEDDKRALPDFN--TFPRIG-AS	356
DB	313	AASPEHRETRARLARGLWNTKEOYLITGRMG--EYDNGLRWQDAPHIRWAGGEVS	370
QY	357	YPSYSAVWYVTLQLRWGMGMINFEPKPNWYLDTPAKVYRFDIYLAAKELVAEGKAAEDP	416

Db 371 FFWLSDGMWFLTQFRRMGLLKQ-APD--YLAVASRINRIDWQAAQ---AVGISA---- 421

QY 417 PADTSIKPSONFIDKVPDPANKENDYLAKEAIGK 453

Db 422 PA--ARMRSSTLMGDTWNGSDPEGYARHFSIORKG 455

RESULT 2

US-09-540-236-2892

; Sequence 2892, Application US/09540236

; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540.236

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 2892

; LENGTH: 474

; TYPE: PRT

; ORGANISM: M.catarhalis

US-09-540-236-2892

Query Match 18.3%; Score 446.5; DB 4; Length 474;

Best Local Similarity 27.9%; Pred. No. 4.5e-36;

Matches 199; Conservative 82; Mismatches 180; Indels 71; Gaps 11;

QY 2 KTIIRSSSKLLTLLTSASLAWGLTAPDVGAVGKL-EKEDLKFGFIKLTDMAPLAVAA 59

Db 79 KTVGGTTA-----LAAIASVLPRIATLQEAALIDTLKPEKSSVDIGFILTATPLIMD 132

QY 60 EKGFFEDGELFVQLEQAAMKVMMDRVANGELDGSMTLAPAPLAASVGGT-KADIEVFF 118

Db 133 PLGYVAEQGKANLKRAGALVRDQMMRELDAAHFLAPMLAINLGSAKQNNKVA 192

QY 119 SMGFNGNAITVSNEIWMQKNPILEGSKVHPKADYLPKVVKEYKAEGKPFNMAMTFP 178

Db 193 IONTNGQALVMA-----LKHKNRNPKNKWKMTFAPLPE 226

QY 179 AGSHNIKRLWYLAAGINPGYSPPODISQIGADALLSVTPPQMPSTLEAGTIFGYCV 238

Db 227 HSHHNYLRLRYFLAEHGLDP-----DKDVYLRLLTTPPDMIAMLKAGNIDGFFG 273

QY 239 GEPMNOQAFYKGIQVIVITDELMKDTPEKVGVTQMAEKYNTYLAATKALIRAILM 298

Db 274 PEPFNORAVWDKAGYIHTLSRDIWNGHPCCSFSTQSFINDPQTLAVYRALIKANVM- 332

QY 299 DADNNKRRKEALEMLAQOQYVADVEVLAASNMGTFEYEKDDKRALPDFTFFRHGAS-Y 357

Db 333 -ANKPSIRKDSLSPAYLNPVELVLOSINGRADGVDITIQDVPD-----RMGFDWM 386

QY 358 PSYSSAVWYLTQLRMGMINFEKPDWYLDATAKNVPRDIYLAALAEVAEGKAKADEFP 417

Db 387 PMHSVAAMMTQMKRWGYIT---GNINYODIANQVF---MLTAKQOMAGYTVADDBP 440

QY 418 ADTSIRKSONFFIDKVPDPANKENDYLAKEAIGLKQRYVAG 459

Db 441 KTVWGKQ-----FNATPDYALDSFALGHKSTGRILHG 473

RESULT 3

US-09-328-352-7572

; Sequence 7572, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328.352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 7572

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7572

Query Match 13.9%; Score 338.5; DB 4; Length 336;

Best Local Similarity 26.5%; Pred. No. 2e-25;

Matches 96; Conservative 63; Mismatches 144; Indels 59; Gaps 9;

QY 34 VCKLEKEDLKFGFIKLTDMAPLAVAAEKGFEDGELFVQLEQAAMKVMMDRVANGELDG 93

Db 1 MSKLEKQDQGYITFLDLOIALMLMAKQGFEEVDLDVTLVKEASWASLRDLARGLDA 60

QY 94 SMMLAPAPLAASVGGTAD-----IEVPSGFGNNAITVSNELWMQ---KRNIPLE 144

Db 61 AHCLSAMLPAAAMG---ADQIGALQTPPLVLSKKNRAFSISQKLIHOLAIRENDNAQTT 116

QY 145 GGGPVHPKADYLPKVVKEYKAEGKPFNMAMTFPAGSHNIKRLRYWLAAGINPGYSPQ 204

Db 117 AOKVIOYIEODH-----TSLAHVFKHSIHYYCLREWLALA----- 152

QY 205 DISQIGADALLSVTPPQMPSTLEAGTIFGYCVGEPMNOQAVFGIGVPIITDELMKD 264

Db 153 --DSRLAQTLKALKALPPRYWMEALDNHVIDGFCVGEPMNTOGELLGSLIVCSQDIIPN 210

QY 265 TREKFGVTQMAEKYNTYLAATKALIRAILWLDADNNKRRKEALEMLAQ---KQYVG 320

Db 211 VADKTLAVTQEWAEQHPHTLVALTTRAIKMAQ--KELSLMKQFAPYLKLVFEGIVRFHCS 268

QY 321 ADVEVLAASNMGTFEYEKDDKRALPDFTFFRHGASYSYSSAVWYLTQLRMGMINPEK 380

Db 269 EEVNDKTYMTIONIKHLVKENAHPQEDFH-----WLFQOMQMKEL-QLA 314

QY 381 PD 382

Db 315 PD 316

RESULT 4

US-09-252-991A-30501

; Sequence 30501, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252.991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074.788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094.190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30501

; LENGTH: 249

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30501

Query Match 12.0%; Score 292.5; DB 4; Length 249;

Best Local Similarity 31.3%; Pred. No. 5.4e-21;

Matches 73; Conservative 40; Mismatches 105; Indels 15; Gaps 6;

QY 218 VTPPQMPSTLEAGTIFGYCVGEPMNOQAVFGIGVPIITDELMKDTPEKFGVTQMA 277

Db 19 VVPRQWGMGHQAGRIDGFCAGPAGALAVDQGGFTATSGAIVPDHREKLTGTTAAV 78

QY 278 EKYPTTYLAATKALIRAILWLDADNNKRRKEALEMLAQOQYVADVEVLAASNMGTFEYE 337

Db 79 DAYPNTABALVMAVVDASRFTL-QNAENRLGTAQLISGRDYVDALGAIQPRFFGRYDQG 137

```

QY 338 KODKRALPDENFFRNG-ASYPSYSAVWLTOLRRMGINERKPDNWTIDTAKNYRPD 396
DB 138 LGAMODPHRLRYADGEVVRPPLSGMMFMTOFRRWGLIRE-DPI--YLGARRVQQA 194
QY 397 IYLAAKEVVAEGKAKADEPPADTSIKPSQNFIDKVPDPANKENDYLAKFAI 449
DB 195 LVRDAATLGL--RLDGADMRST-----LIDRTWDGSDPAGYASFPFI 237

RESULT 5
US-09-252-991A-25707
; Sequence 25707, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25707
; LENGTH: 419
; TYPE: PR
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25707

Query Match
Best Local Similarity 24.6%; Score 240; DB 4; Length 419;
Matches 93; Conservative 68; Mismatches 141; Indels 76; Gaps 14;

QY 9 SKKLLTSLASLAVWG-LTAPDVGAVKLEKE-DKFGFKLTDMAPIAVALAEKGFED 66
DB 29 SRRDIKTLALSAAGALPLSLQRAAPAPRIGLPTDTPPLVAVANGLFPA 88
QY 67 EGLFVQLEAQA--NMKVMDRVVNGELDGSMLAPAPLAASFGSTKADIEVPSMGFN 123
DB 89 EG--IOAERPVLLRSNAQVIEAFISQGVAVIHLSMTYWAR--YSSKVPATYVANNHYG 144
QY 124 GNAITVSENIWHQMKENIPLEGKRVNPIKADYLKPVVEKYAEKGFNMAMTFPAGSHN 183
DB 145 GGLGLVAPRI-----ADVRLGGKSV-----AIPFWSYSHN 175
QY 184 IKLRVYLAAGINPGYSPPODSGOIGA-----DALLSVTPPPQMPSTLEAGTIFGCV 238
DB 176 VVLQQLLRDNGL-----RAVSRAVGAALADEVNLVLLPSPDMPPLASRKHGYIV 227
QY 239 GEPANQOAVFKGIVGITDEELMKDTPKVFQVTQMAEKYENTYLAATKALIRAIWL 298
DB 228 AEPFNLALBNLKVGRVORTGDMWRNHACCVCVPMHEDERRPQWQKYNVAIVKQQL- 286
QY 299 DADNNKRKEALEMTAQ--KQYGVADVEVL-----AASMGSTFEYKQDKKA 343
DB 287 ---TEHRAEAQAQLSKAGANRTPAPAEVLGVLAPGAEEQAYLASAGAIRHADQERR 343
QY 344 LPDENFFRNGASYPSS 361
DB 344 I-DFOPY-----PYPSYT 355

RESULT 6
US-09-602-787A-384
; Sequence 384, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Mark
; APPLICANT: Krüger, Burkhard
; APPLICANT: Schöder, Hartwig

```

```

; APPLICANT: Zelder, Oskar
; APPLICANT: Habener, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602.787A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5

```

```

: PRIOR FILING DATE: 1999-09-03
: PRIOR APPLICATION NUMBER: DE 19942079.3
: PRIOR FILING DATE: 1999-09-03
: PRIOR APPLICATION NUMBER: DE 19942088.2
: PRIOR FILING DATE: 1999-09-03
: NUMBER OF SEQ ID NOS: 678
: SEQ ID NO 384
: LENGTH: 275
: TYPE: prt
: ORGANISM: Corynebacterium glutamicum
US-03-602-787A-384

```

Query Match	9.0%;	Score 219.5;	DB 4;	Length 275;
Best Local Similarity	23.1%;	Pred. No. 1.5e-13;		
Matches 66;	Conservative 52;	Mismatches 127;	Indels 41;	Gaps 5

Qy 42 KKFSGITLMDPAVLVAARKFFEEBGLVTOLEAOANMVKYMDRVNNGEISDSSHTALAP 101
Db 26 LTTGVPPLIAGSAPIALADALGLEFKKHGVNLTAKTISGMSDLTAVATTEQJLDVNMHTSPMT 85
Qy 102 LAASVGF-GTKADIEVPEFSGFNAGAITVSNELMHWQMKENIPLEGKPVHPKADYIKPV 160
Db 86 VAINAGVTNARSPTLSFTQNTNGQITLASKHYGSVNSAALDKG----- 130
Qy 161 VEKYAEKGFENMAMTFPAGSHNIKLRVYLAAGINPGYYSPPQDISQIGADALLSYTP 220
Db 131 ---MVLGIPFEYSV-----HALLRLDPLVNSAADP-----IADLEIRLLR 167
Qy 221 PPMQESTLEAGTIFGYCVSGEPNNQOAVFGIIGVPTTDEELMKDTEPEKVGTVQKMAEKY 280
Db 168 PADWAOGLTVEGIDEGTIGCPFNERALISNGSGRIWLTTLQMDMKHCCAVAMAKEMKAHE 227
Qy 261 PNTYLAVTKALIRALAIWLDADNNKKNKEAIEMLAQKQYQYGADEVLT 326
Db 228 PTAAGVNAALAEASAIL--SNPAQDSSARTLSCEKRYINOPATLL 271

```

RESULT 7
US-09-252-991A-30392
; Sequence 30392, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS.
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30392
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30392

```

Query Match	6.4%;	Score 157;	DB 4;	Length 197;
Best Local Similarity	38.9%;	Pred. No. 1.7e-07;		
Matches 37;	Conservative 14;	Mismatches 42;	Indels 2;	Gaps 1;

QY 38 EKEDEIKFGFIKLTDMAPLVAAEKGFEEDEGLFVLELAQANKVMVMDRVYNGEIDDSHML 97
Db 78 EKSLAIDIGFMAITDASLVAATQGFAPQYGLTNRROPSPWATLKDILSGELDAQCL 137
QY 98 APAPLAASVGFQTKADIEVPFSMGF--NGNAITVS 130
Db 138 YGLVYGVOLIGLGSAASEMAVLMGLCONCAINLS 172

RESULT 8

```

US-09-902-540-12264
; Sequence 12264, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goleman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10 (15849) B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12264
;
; LENGTH: 336
;
; TYPE: PRT
;
; ORGANISM: Myxococcus xanthus
US-09-902-540-12264

```

Query Match	5.0%;	Score 121.5;	DB 4;	Length 336;
Best Local Similarity	23.8%;	Pred. No. 0.0016;		
Matches	73;	Conservative	35;	Mismatches 116;
				Indels 83;
				Gaps 12

```

QY 12 LLLTSLASLAWGLTIAPDVGAVGKLEKEDLKEGFIKLTMDMPLVAABEGFEDEGLFV 71
Db 7 LIPVVVAGIATVAGACCKREPSAARG--PDAFLKLGFFPNNTTHAOLVGNAGETIASOQVG 64
QY 72 QLEA-QAN-WKVYMDRVVNGELDGSHTLAPAPLAASVGFQ-----TKADIEVPFSWG-F 122
Db 65 RLEVVQFNAGPAAAMEALVAGSLDVSF-----VSGSPAINTEFLKAGRELRIVAGAV 114
QY 123 NGNATTVSNEIMHQKKPNIPLEGKRPVHPKADYLKPVVEKKAKBEKPFNMATPFAGSH 162
Db 115 NNGAVLVVRTV-----KTPAEIKGKKGLASPOLGNTQ 145
QY 163 NIKLRYMLAAGINQGYSPDIDISGQIGADALLSVTPRPQMPSTLEAGTIFGYCVGEPM 242
Db 146 DIALRYMLKQOGL-----TTHDGGDV---QIFPLSNPDIIIGQFIRGC-IEGAWBEPM 196
QY 243 NOQAAFKGIGVIVTDEELMKDTPKEKVFVETQMAEKAYNTYLAVTK-----AL 291
Db 197 CARLVAEKGKRLVNEKDIMP-----GGRFFTYLVTTROYLETORPRVALL 243
QY 292 IRAAIWL 298
Db 244 LRAHVRL 250

```

```

: RESULT 9
: US-09-252-991A-26757
: Sequence 26757, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 26757
: LENGTH: 337
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-26757

```

Query Match 4.8%; Score 116.5; DB 4; Length 337;


```
/ ADDRESS: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD-ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/
/ SOFTWARE: ASCII
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 5754:
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 369 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/
/ FEATURE:
/ NAME/KEY: m1sc.feature
/ LOCATION: (8) LOCATION 1...369
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5754:
US-09-107-532A-5754

Query Match      4.2%; Score 102; DB 4; Length 369;
Best Local Similarity 21.5%; Pred. No. 0.17;
Matches 85; Conservative 47; Mismatches 138; Indels 126; Gaps 20;

QY 10 KTLTLTSLASLAWGLTIAPDVGA---VGKLEKEDLKFGFIKLTMDAPLAVAAGKGFED 66
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 41 KKLWLLPLLLLSACGTAKETSSKQEKIDKLKVTLLVDYVPNTNHTGIYLAKEKGYRKE 100
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 67 EGLFVOL-----EAQANKKVVMDRVVNGELDGSHTMLAPAPLAASY----- 106
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 101 AGLVNQLIERGDNSTIGLVGADKQPGVSYQED-VTYAHADQGN--FVKAIATYIKIN 157
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 107 --GFGTKADIEVPFSGNGENALITVSNEIMHQMKPNIPLGGKRVHPRIKADYLYKPVVEKY 164
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 158 TSGFATLSDSNHISPDGFKTYAG---WQS-----PSEEAIVLKAVMEK- 198
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 165 KAGKRFNNAMTFPPAGSHNIIKRYWLAAGINPGYSPPODISGQ--GADALLSTVPPQ 223
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 199 --DGGDFS-----KLT-MVSGNGEGESLIGKSSDIQWYFEGWD----- 233
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 224 MESTLAGTIFGCVGEPNNQAVFKGIGVPI-ITDEILMKDTPPEKVFVGTKQMAEKYPN 282
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 234 MIKAKAGGLEINITYPLKEIDERLDY---TPVITTDQLIKSPFLV-----Q 278
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 283 TYLAIVTKALIRAAIWLADNNKRRKEAIEMLAQKQYVGADVEVLAAMNGTFEYKDDKR 342
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 279 SFWDATKKGYQEAR--KDPNDSAK-----LLOK-----YAKENDR 311
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 343 ALPDFNTFPRHGAASYPSYSAWYILTQLRMGMINE 378
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 312 -----FTLEESQAFLSKN-----YTDDPRNWGIMEE 337
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```
RESULT 13
US-09-397-550-4
/ Sequence 4, Application US/09397550
/ Patent No. 6783952
/ GENERAL INFORMATION:
/ APPLICANT: Warner-Lambert
/ TITLE OF INVENTION: Secreted soluble alpha2 delta-2, alpha2 delta-3, alpha2
/ TITLE OF INVENTION: delta-4 calcium channel subunit polypeptides and
/ TITLE OF INVENTION: screening assays using same
/ FILE REFERENCE: 180
/ CURRENT APPLICATION NUMBER: US/09/397,550
/ CURRENT FILING DATE: 1999-09-16
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 1062
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-397-550-4

Query Match      4.2%; Score 102; DB 4; Length 1062;
Best Local Similarity 23.7%; Pred. No. 0.95;
Matches 79; Conservative 42; Mismatches 142; Indels 70; Gaps 17;

QY 7 SSKKLLTSLASLAWGLTIAPDVGAAGVGLKLEKEDLKFGFIKLTMDAPLAVAAGKGFED 66
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 287 SSPKDMVILVDVSGSVSLTL-----KLMKTSVCEMLDPTLSDDDVYNVASFPEKQAP 338
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 67 EGLFVOL-EAQANKKVVMDRVVNGELDGSHTMLAPAPLAASYVGFGTKADIEVPFSGNGEN 125
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 339 VSCFTHLVQANVRNKKVFKEAVQ-----MVAKGTTGYKAGFEVAFDQLQNSN 386
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 126 AITVSNEIMHQMKPNIPL--EGKRVHPRIKADYLYKPVVEKYKAEKPRNMAMTFPPAGSHN 183
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 387 -ITRAN-----CNKIMIMFTDGE-----DRVQDVEKYNWPNRTVRV-FTFSVGOHN 432
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 184 IKLR--YWLAAGINPGYSPPODISG-QIGADALLSTVPPOMPSTLEAGTIFGVCVGE 240
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 433 YDVTLQMMACA--NKGYTFEIPISGALRINTQETLVDLGRMVLAAGEAKYV----- 483
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 241 PNNQAVRK-GIGVEVITDEILMKDTPPEKVFVGTKQMAEKYPNTYLAIVTKALIRAAIWD 299
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 484 QWTN--VYEDALGLGLVVTGTL-----PVFNLTQDGGGEKKN-----QLILGVMGID 528
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 300 ADNNKRRKEAIEMLAQKQYVGADVEVLAAMNG 332
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 529 VALND-----IKRLTPNNTLGGANGVFAIDING 556
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14
US-09-470-443-6
/ Sequence 6, Application US/09470443
/ Patent No. 6441156
/ GENERAL INFORMATION:
/ APPLICANT: Lerman, Michael I.
/ APPLICANT: Manna, John D.
/ APPLICANT: Latif, Farida
/ APPLICANT: Wei, Ming-Hui
/ APPLICANT: Sekido, Yoshitaka
/ APPLICANT: Duh, Fuh-Mei
/ TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
/ FILE REFERENCE: NIH-05043
/ CURRENT APPLICATION NUMBER: US/09/470,443
/ CURRENT FILING DATE: 1999-12-22
/ EARLIER APPLICATION NUMBER: 60/114,359
/ EARLIER FILING DATE: 1998-12-30
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 1076
/ TYPE: PRT
```


ORGANISM: Homo sapiens
US-09-470-443-6

Query Match 4.2%; Score 102; DB 4; Length 1076;
Best Local Similarity 23.7%; Pred. No. 0.97;
Matches 79; Conservative 42; Mismatches 142; Indels 70; Gaps 17;

```
QY 7 SSSKLLLTLSASLAVNGLTIPADVGAVGLEKEDLKFGFIKLTDMAPLVAAEKGFED 66
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 218 SSRKDMVITVDVSGSVGLTL-----KLMKTSVCEMLDTSDDYVNVASFNEKAQP 269
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 67 EGLFVQL-EAQAAMKVMMDRVVNGELDGHMLAPAPLAASVGFCTKADIEVPSMGFNGN 125
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 270 VSCFTHLVQANVANKKVFKEAVG-----MVAKGTGYKAGFEYAFDQLQNSN 317
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 126 AITVSEIWHQMKPNIPL--EGGKPVHPKADYIKPVVEKYKAEGKPFNNAMTFPAGSHN 183
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 318 -ITRAN-----CNKMMIMFTDGE-----DRVQDFEKKYNNPNTVRV-FTFSVQGHN 363
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 184 IKLR--YMLAAGINPGYSPPODISG-QIGADALLSVTPPPQMPSTLEAGTIFGYCVGE 240
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 364 YDVTPLQWMAKA--NKGYYEIPISGAIKINTQEIYDLVGRPMVLAGEKAQV----- 414
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 241 PMNQAVFK-GIGVPVITDELMKDFPEKVFYTKQWAEKYPNTYLAVTKALIRAAIWL 299
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 415 QWTN--VYEDALGLGLVVTGTL-----PVFNLTQDGPGKEKN-----QLILGWGID 459
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 300 ADNKKRKEAIEMLAQKQYVGVADVEVLAASMG 332
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 460 VALND-----IKRLTPNYTLGANGVYFAIDLNG 487
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 15

US-09-397-550-5
; Sequence 5, Application US/09397550
; Patent No. 6783952
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert
; TITLE OF INVENTION: Secreted soluble alpha2 delta-2, alpha2 delta-3, alpha2
; TITLE OF INVENTION: delta-4 calcium channel subunit polypeptides and
; FILE REFERENCE: 180
; CURRENT APPLICATION NUMBER: US/09/397,550
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1082
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-397-550-5

Query Match 4.2%; Score 102; DB 4; Length 1082;
Best Local Similarity 23.7%; Pred. No. 0.98;
Matches 79; Conservative 42; Mismatches 142; Indels 70; Gaps 17;

```
QY 7 SSSKLLLTLSASLAVNGLTIPADVGAVGLEKEDLKFGFIKLTDMAPLVAAEKGFED 66
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 287 SSPKDMVITVDVSGSVGLTL-----KLMKTSVCEMLDTSDDYVNVASFNEKAQP 338
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 67 EGLFVQL-EAQAAMKVMMDRVVNGELDGHMLAPAPLAASVGFCTKADIEVPSMGFNGN 125
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 339 VSCFTHLVQANVANKKVFKEAVG-----MVAKGTGYKAGFEYAFDQLQNSN 386
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 126 AITVSEIWHQMKPNIPL--EGGKPVHPKADYIKPVVEKYKAEGKPFNNAMTFPAGSHN 183
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 387 -ITRAN-----CNKMMIMFTDGE-----DRVQDFEKKYNNPNTVRV-FTFSVQGHN 432
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 184 IKLR--YMLAAGINPGYSPPODISG-QIGADALLSVTPPPQMPSTLEAGTIFGYCVGE 240
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 433 YDVTPLQWMAKA--NKGYYEIPISGAIKINTQEIYDLVGRPMVLAGEKAQV----- 483
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 241 PMNQAVFK-GIGVPVITDELMKDFPEKVFYTKQWAEKYPNTYLAVTKALIRAAIWL 299
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```
DB 484 QWTN--VYEDALGLGLVVTGTL-----PVFNLTQDGPGKEKN-----QLILGWGID 528
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 300 ADNKKRKEAIEMLAQKQYVGVADVEVLAASMG 332
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 529 VALND-----IKRLTPNYTLGANGVYFAIDLNG 556
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Search completed: September 16, 2005, 00:49:10
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 00:47:41 ; Search time 172 Seconds
(without alignments)
1092.363 Million cell updates/sec

Title: US-10-689-200-2

Perfect score: 2443

Sequence: 1 MKTIRSSSKLLTLTSLA.....AKFAIGLKQIVAGKAVD 464

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2443	100.0	464	US-10-689-200-2	Sequence 2, Appl1
2	525	21.5	403	US-10-282-122A-69975	Sequence 69975, A
3	240	9.8	399	US-10-282-122A-66370	Sequence 66370, A
4	240	9.8	399	US-10-389-647-502	Sequence 502, App
5	234.5	9.6	396	US-10-282-122A-49807	Sequence 49807, A
6	227	9.3	257	US-10-282-122A-44527	Sequence 44527, A
7	226	9.3	294	US-09-738-626-4906	Sequence 4906, App
8	222.5	9.1	372	US-10-282-122A-68363	Sequence 68363, A
9	220.5	9.0	391	US-10-282-122A-49957	Sequence 49957, A
10	219.5	9.0	275	US-10-627-476-384	Sequence 384, App
11	217.5	8.9	391	US-10-282-122A-48182	Sequence 48182, A

12	144	5.9	342	US-10-282-122A-72603	Sequence 72603, A
13	140	5.7	192	US-10-282-122A-47676	Sequence 47676, A
14	116.5	4.8	883	US-10-156-761-8511	Sequence 8511, App
15	113	4.6	1355	US-10-450-763-36085	Sequence 36085, A
16	111.5	4.6	353	US-10-156-761-9969	Sequence 9969, App
17	111	4.5	1357	US-10-450-763-39994	Sequence 39994, A
18	108	4.4	533	US-10-425-118-51223	Sequence 51223, A
19	107	4.4	518	US-10-424-599-198653	Sequence 198653, A
20	106.5	4.4	461	US-10-282-122A-46659	Sequence 46659, A
21	105	4.3	419	US-10-282-122A-46437	Sequence 46437, A
22	105	4.3	471	US-10-425-115-185585	Sequence 185585, A
23	105	4.3	1390	US-10-275-595A-6	Sequence 6, Appl1
24	105	4.3	3352	US-10-156-761-7861	Sequence 7861, App
25	105	4.3	4435	US-10-287-228-304	Sequence 304, App
26	104	4.3	866	US-10-241-596-104	Sequence 104, App
27	104	4.3	1420	US-10-241-596-110	Sequence 110, App
28	103.5	4.2	280	US-09-815-242-11714	Sequence 11714, A
29	103	4.2	864	US-10-241-596-102	Sequence 102, App
30	103	4.2	1139	US-10-417-375-176	Sequence 176, App
31	103	4.2	1141	US-10-417-375-174	Sequence 174, App
32	103	4.2	1142	US-10-417-375-172	Sequence 172, App
33	102.5	4.2	320	US-10-282-122A-55774	Sequence 55774, A
34	102	4.2	613	US-10-369-493-18239	Sequence 18239, A
35	102	4.2	860	US-10-241-596-175	Sequence 175, App
36	102	4.2	862	US-10-241-596-94	Sequence 94, Appl
37	102	4.2	866	US-10-241-596-88	Sequence 88, Appl
38	102	4.2	867	US-10-241-596-96	Sequence 96, Appl
39	102	4.2	867	US-10-241-596-98	Sequence 98, Appl
40	102	4.2	870	US-10-241-596-92	Sequence 92, Appl
41	102	4.2	871	US-10-241-596-84	Sequence 84, Appl
42	102	4.2	871	US-10-241-596-86	Sequence 86, Appl
43	102	4.2	871	US-10-241-596-90	Sequence 90, Appl
44	102	4.2	1062	US-10-902-531-4	Sequence 4, Appl1
45	102	4.2	1076	US-10-116-949-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-10-689-200-2
Sequence 2, Application US/10689200
Publication No. US20040126848A1
GENERAL INFORMATION:
APPLICANT: DiCosimo, Deana J.
APPLICANT: Ni, Hao
APPLICANT: Ye, Rick
APPLICANT: Picataggio, Stephen
APPLICANT: Wang, Tao
TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN CL METABOLIZING BACTERI
FILE REFERENCE: C11747 US NA
CURRENT APPLICATION NUMBER: US/10/689,200
CURRENT FILING DATE: 2003-10-20
PRIOR APPLICATION NUMBER: 60/419,872
PRIOR FILING DATE: 2002-10-21
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 464
TYPE: PRT
ORGANISM: Methylobacillus sp. 16a
US-10-689-200-2

Query Match 100.0%; Score 2443; DB 16; Length 464;
Best Local Similarity 100.0%; Pred. No. 3.2e-216; Indels 0; Gaps 0;
Matches 464; Conservative 0; Mismatches 0;

QY 1 MKTIRSSSKLLTLTSLAVWGLTIAPDVGAAGKLEKEDLKFGLTDMAPLAAVAE 60
DB 1 MKTIRSSSKLLTLTSLAVWGLTIAPDVGAAGKLEKEDLKFGLTDMAPLAAVAE 60
QY 61 KGFEDDEGLFVQLEAQAAMKVMRVRVNGELDGSHMLAPAPLAASVGFQTADEVPSFM 120

```
Db      61 KGFEEDEGLFVQLEAQANKVMDRVVNGELDGSHTLADAPLAASVGFCTKADIEVPSM 120
Qy      121 GENGNAITVSNETWHQMKNPIPLEGGKPVHPKADYLVKVEKYKAEKPFVMAWTFPAG 180
Db      121 GENGNAITVSNETWHQMKNPIPLEGGKPVHPKADYLVKVEKYKAEKPFVMAWTFPAG 180
Qy      181 SHNIKLRVYLAAGGIPGYSPPODISGQIGADALSTVPPOMPSTLEAGTIFGCYGE 240
Db      181 SHNIKLRVYLAAGGIPGYSPPODISGQIGADALSTVPPOMPSTLEAGTIFGCYGE 240
Qy      241 PMNQAVFEGKIGVPTTDELMKDPTEKVFVGTQKQAEKYPNTYLAVTKALIRAAITWDA 300
Db      241 PMNQAVFEGKIGVPTTDELMKDPTEKVFVGTQKQAEKYPNTYLAVTKALIRAAITWDA 300
Qy      301 DNNKPKKEAIEMLAQKQYGVADVEYLAASNGTFFEEKDKRALPDPNTPFFHGSYSBY 360
Db      301 DNNKPKKEAIEMLAQKQYGVADVEYLAASNGTFFEEKDKRALPDPNTPFFHGSYSBY 360
Qy      361 SSADVWLTQLRRGMNTEPKPDNMYLDTAKNYRPPITYLAAAKEIVAEKAKADEPPADT 420
Db      361 SSADVWLTQLRRGMNTEPKPDNMYLDTAKNYRPPITYLAAAKEIVAEKAKADEPPADT 420
Qy      421 STKPSQNFPIDKVPFDANKPNDYLAKEFALGLKQKQTVAGAKYVD 464
Db      421 STKPSQNFPIDKVPFDANKPNDYLAKEFALGLKQKQTVAGAKYVD 464
```

```
RESULT 2
US-10-282-122A-69975
; Sequence 69975, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69975
```

```
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69975
```

```
Query Match 21.5%; Score 525; DB 15; Length 403;
Best Local Similarity 31.7%; Pred. No. 2,5e-39;
Matches 132; Conservative 72; Mismatches 164; Indels 48; Gaps 12;
```

```
Qy      38 EKEDKRFQFILTDMAPLVAAEKGFEEDEGLFVQLEAQANKVMDRVVNGELDGSHTL 97
Db      19 ENNSTLDVGFPMALTDAPLVVAATQGFAPYGLSTLTKQTSWAGLRDLVSGQLAAHSL 78
Qy      98 AAPALAASVGF--GTRKADIEVPSWGFNGNAITVSNETWHQMKNPIPLEGGKPVHPKAD 155
Db      79 YGLIYAVELGISGGPATMALLMGLNQGCCINLSRF-----LQDAGVITPEALD 128
Qy      156 YLKPVEKYKAEKGFNNAMTFPAGSHNIKLRVYLAAGGIPGYSPPODISGQIGADAL 215
Db      129 -----KRAHQSGSKLTFAGQTFPTGNHAMWLYWMLASQGIHP-----LDDV 168
Qy      216 LS-VTPFPOMPSTLEAGTIFGCYGEFPWNOQAVFEGKIGVPTTDELMKDPTEKVFVGT 274
Db      169 TSVVVPPTQMAOHLQAGRIDGFCVGEFPWASAVOQDLFTWATSGAIVPDHKGVLGCTR 228
Qy      275 QWAEKYPNTYLAVTKALIRAAITWLDADNNKPKKEAIEMLAQKQYGVADVEYLAASNGTFF 334
Db      229 EPEQNPNTARALLMAVLEASRFIE-QSDHRRRSTQQLSVDYDIALDCTIEPRLIQY 287
Qy      335 EYKDKRALPDPNTPFFHGS--ASYPSSSAVWYLTQLRRGMNTEPKPDNMYLDTAKNY 393
Db      288 SDGLGNQMDPHAVSFHQGVNYPWLSGMWFMFTQFRWGLIRE-DDD--YLAVASRVQ 344
Qy      394 RPDYLAAKEIVAEKAKADEPPADTSTKPSQNFPIDKVPFDANKPNDYLAKEFAL 449
Db      345 QLDLYRQANAL-----GIDAPEAT-LRSSQ--LIDGVKWDGSDPAGYARSFYL 390
```

```
RESULT 3
US-10-282-122A-66370
; Sequence 66370, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
```

PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 66370
LENGTH: 399
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66370

Query Match 9.8%; Score 240; DB 15; Length 399;
Best Local Similarity 24.6%; Pred. No. 4.9e-13;
Matches 93; Conservative 68; Mismatches 141; Indels 76; Gaps 14;

QY 9 SKLLLTLSASLAWG-LTTPADVGVKLEKE-DLKFETIKLTDMAPLVAAEKGFED 66
DB 9 SRDILKLAALLSAAGALPLSLQRAAEPDAPRIGYLPITDAPLIVAHANGLFPA 68
QY 67 EGLFVQLEQA---NMKVMDRVNGEELDGSMLAPAPLAASVGFETKADIEVPSMGFN 123
DB 69 EG--IQAERPVLIRSWAQVTEAFISGVNVHLSPTVWAR--YGSKPAPAKVANNHYG 124
QY 124 GNAITVSNIEIWMQKNIPLEGKPVHPKADYLKVEKRYKAEGKPFMMATPFASHN 183
DB 125 GSGLTVAPEI-----ADVQLOGKSV-----AIPFWWSIHN 155
QY 184 IKLRVWLAAGINPGYSPPODISGOIGA-----DALSTPPPMPTLEAGTIGYCV 238
DB 156 VVLQQLLRDNGL-----RAVSRVAGALAADENVLVLPPSDMPPALASKRIRIGYIV 207
QY 239 GEPNNOQAVFKGIGVPTIDBELMKDTPKRVGVTQMAEKYNTYLAATKALIRAIWL 298
DB 208 AEPFNALAEMLKVGVRFTGVDWRNHACCVMFHEHDLERRPQWSQKVNAIVKQOLM- 266
QY 299 DADNNKREKALEMLAQ---KQYGVADVEVL-----AASMGTFEYKDDKRA 343
DB 267 ---TRHRAEAQQLSKAGANRTTPHAPVGLGVLAPGAEOQAYLASGAIIRADWQERR 323
QY 344 LPDNTFFRHGASYPSSYS 361
DB 324 I-DFQPY-----PYPSYT 335

RESULT 4
US-10-389-647-502
Sequence 502, Application US/10389647
Publication No. US2004003349A1
GENERAL INFORMATION:
APPLICANT: GREENBERG, E. Peter
APPLICANT: SCHUSTER, Martin
APPLICANT: LOSTROH, Candi
TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
FILE REFERENCE: UZ-038CP
CURRENT APPLICATION NUMBER: US/10/389,647
PRIOR FILING DATE: 2003-03-14
PRIOR APPLICATION NUMBER: 09/653730
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/153022
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 710
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 502
LENGTH: 399
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-389-647-502

Query Match 9.8%; Score 240; DB 15; Length 399;

Best Local Similarity 24.6%; Pred. No. 4.9e-13;
Matches 93; Conservative 68; Mismatches 141; Indels 76; Gaps 14;

QY 9 SKLLLTLSASLAWG-LTTPADVGVKLEKE-DLKFETIKLTDMAPLVAAEKGFED 66
DB 9 SRDILKLAALLSAAGALPLSLQRAAEPDAPRIGYLPITDAPLIVAHANGLFPA 68
QY 67 EGLFVQLEQA---NMKVMDRVNGEELDGSMLAPAPLAASVGFETKADIEVPSMGFN 123
DB 69 EG--IQAERPVLIRSWAQVTEAFISGVNVHLSPTVWAR--YGSKPAPAKVANNHYG 124
QY 124 GNAITVSNIEIWMQKNIPLEGKPVHPKADYLKVEKRYKAEGKPFMMATPFASHN 183
DB 125 GSGLTVAPEI-----ADVQLOGKSV-----AIPFWWSIHN 155
QY 184 IKLRVWLAAGINPGYSPPODISGOIGA-----DALSTPPPMPTLEAGTIGYCV 238
DB 156 VVLQQLLRDNGL-----RAVSRVAGALAADENVLVLPPSDMPPALASKRIRIGYIV 207
QY 239 GEPNNOQAVFKGIGVPTIDBELMKDTPKRVGVTQMAEKYNTYLAATKALIRAIWL 298
DB 208 AEPFNALAEMLKVGVRFTGVDWRNHACCVMFHEHDLERRPQWSQKVNAIVKQOLM- 266
QY 299 DADNNKREKALEMLAQ---KQYGVADVEVL-----AASMGTFEYKDDKRA 343
DB 267 ---TRHRAEAQQLSKAGANRTTPHAPVGLGVLAPGAEOQAYLASGAIIRADWQERR 323
QY 344 LPDNTFFRHGASYPSSYS 361
DB 324 I-DFQPY-----PYPSYT 335

RESULT 5
US-10-282-122A-49807
Sequence 49807, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Cair, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16

```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49807
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-282-122A-49807

Query Match
Best Local Similarity 23.3%; Pred. No. 1,6e-12;
Matches 105; Conservative 77; Mismatches 150; Indels 119; Gaps 18;

9.6%; Score 234.5; DB 15; Length 396;

9 SKLLTLTSLASLAWGTLTAPDVGAVGKLEKED---LKEFGIKLTDMAPLAAGAAGFF 64
DB 7 SRREWKLASMTFVAG--AAPLLAALNARAENPDAPRIGYLPITDAAPLVAHNNYF 64
65 EDEGLFVQLEA-QANRVVMDRVVNGELDGSNMLAPPLAASVFGTKADIEVPSMGFN 123
DB 65 DASGLAVEKPTLLRSNAQIVAEFLSGQVNVVHLLAPMTIAR--YGSQAPAKVVAHNNY 122
124 GNAITVSNEMHQMKNRIPLGGKPVHPRIKADYLKVEVEKYKAEGRPMAMTFPAGSHN 183
DB 123 GSALTVAAPDI-----GKLGLGGKTV-----AVPFWYSIH 153
184 IKRLVWLAAGINPGYSPQDISGOIGADAL-LSTVTPPQMPSTLEAGTIFGVCVGEPM 242
DB 154 VVVQVHMLRAGL-----VPLKEKDELKANENVLLIMSBEDMPALASQIAGFIYAEFP 208
243 NQOAVFKGIGVPVITDELMKD-----TPKVFQVTKQMAEKYPNTYLAATKALIRAA 295
DB 209 NAAAEELKVGKVLRTFGDVWKNACCVFWEHERDLTERAASQK-----VVDVAVKQ 261
296 IWLDAQNNKRRKAIEMLA---QKVYGVADVEVLASMGTFEYKEDKRALPDPTFFR 352
DB 262 VWTRRA---HPQEAQDLLSGNNHYTPHSANVLT---VVLADPPDEGTYLAD----- 308
353 HGASYSYSSAVVYLTQLRRWGMINEFKPDNMYLTAKNV-YRPDIYLAALKELY----- 406
DB 309 -----RAITHADNH-----AKRIDPPYPPYATBELVARKLA 341
407 --AEGKAK-----AEDFPADTSIKPS 425
DB 342 TVVEGNAQQLHQLDPAFVARDLVDRFVKKS 372

RESULT 6
US-10-282-122A-44527
; Sequence 44527, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```

; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44527
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; FEATURE: MISC_FEATURE
; NAME/KEY: (7)
; LOCATION: (7)..(7)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-44527

Query Match
Best Local Similarity 9.3%; Score 227; DB 15; Length 257;
Matches 69; Conservative 52; Mismatches 122; Indels 50; Gaps 7;

9.3%; Score 227; DB 15; Length 257;

91 LDGSHMLAPPLAASVFGTKAD-----IEVPSMGFGNATVSENHQW---KPI 141
DB 1 LDAHGXSLMPLAAMG---ADQIGIALQITPLVSKRAFISLQKLYOLAIGESDA 56
142 PLEGGKPVHPRIKADYLKVEVEKYKAEGRPMAMTFPAGSHNIKRLVWLAAGINPGYVS 201
DB 57 QTTAKVLYIYBQD-----TSLAHVFKSIHHYGLCREWMLALA----- 95
202 PPQDISGOIGADALLSTVTPPQMPSTLEAGTIFGVCVGEPMNQAVFKGIGVPVITDELT 261
DB 96 -----DSRIAQTLTKKALPPRYWEALDNHYIDGFCVGEPMNTQBELGLSKIVCSSODI 150
262 WKDTEKVFQVTKQMAEKYPNTYLAATKALIRAAIWLADNNKRRKAIEMLAQKQYVGA 321
DB 151 IPNVADKVLATVQENAEQHPTLVALTAIMKAQ--KEISNKKDRAPIKLK----- 200
322 DVEVLAASNGTFEYKEDKRALPDPTFFRGAASPSYSASVAVYLTQLRRWG 374
DB 201 -VEFGIVRPHGSEVYHVDKYWIGNIVKYLVEKNAAPQREDPHMLFQGMQKMG 252

RESULT 7
US-09-738-626-4906
; Sequence 4906, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
```

```

; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4906
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4906

```

Query Match 9.3%; Score 226; DB 9; Length 294;

Best Local Similarity 23.6%; Pred. No. 6,1e-12; Matches 78; Conservative 56; Mismatches 152; Indels 44; Gaps 6;

```

QY 1 MKTIIRSSSKRLLTLSASLAWGLTAPDVGAVKLEKED---LKEGFIKLTDMAPLAV 57
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MTHILFDSRRFLQGFASLSTLAGAARYVTSTSNNEPADNPTLTIGVPIAGSAPIM 60
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 58 AAKGFEDEBGLVQLEAQNMYVMDRVVNGELDSHMLAPPLASVGF-GTKADIEV 116
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 ADALGFKKHGVVTLTKKYSGMEDLTAYATEQLDVAHMLSPMTVAINGVTNASSRPT 120
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 PFSMGFNGNAITVSNELHMQMKNPILGEGKPVHPKADYLVKVEKYAEKGFEMAMT 176
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 SFQNTNGAHTLASKHYGSVNSAADIKG-----MYLGIPEFYSV- 160
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 FPAASHNIKLRVYLAAGINPGYSPDISQIGADALLSVPPPMESTLEAGTIFGY 236
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 161 ----HALLRDVLVNSAVDP-----IADLELRRLRPDMVAQGLVEGIDGF 202
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 CGEPPNQAVFVGIGVPVITDEELMKDPEKVFQVTKQMAEYENYLAVTALIRAI 296
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 203 IGPFPNERAISNGSGRIWLTQLWDKPCAVAAKEKAEHPRAAQGLNALBEASA 262
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 WLDADNNKRRKEALEMLAKQYGVADVEVL 326
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 263 IL--SNPAQFDSARTLSQEKYLNQPATL 290
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8

```

US-10-282-122A-68363
; Sequence 68363, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68363
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68363

```

Query Match 9.1%; Score 222.5; DB 15; Length 372;

Best Local Similarity 23.1%; Pred. No. 1.8e-11; Matches 90; Conservative 71; Mismatches 139; Indels 89; Gaps 17;

```

QY 42 LKEGFIKLTDMAPLVAAEKGFEEDEGLFVQLEA--QANKVYMDRVVNGELDSHMLA 100
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 17 VAGIVPIIDATPPLVAHNNGLFEAEGIKAEKRPVLLRSQAQYIEAFISQVNVHLSFM 76
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 PLAASVGFCTKADIEVPFSGFNGNAITVSNELHMQMKNPILGEGKPVHPKADYLV 160
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 77 TWAR--YGSKPAKAVANNHVGSGSLTVAAPDI-----SAVKQLGGKTV----- 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 161 VEKYAEKGFEMAMTFPAGSHNIKLRVYLAAGINPGYSPDISQIGADAL--LSVT 219
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 -----ALPFWYSIHNVVYQQLNDNGLTP--VSKP--ANAQLANNEVNLVL 161
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 220 PPMPESTLEAGTIFGYCGEPPNQAVFVGIGVPVITDEELMKD-----TPKEVFGV 272
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 162 PPSDMPPLASGRIGIYIAEFPMALAEHLKXGRVRFQGDVWRHACCVVFMHEHDLAN 221
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 273 TKQMAEKYPNTYLAVTALIRAIWLDADNNKRRKEALEMLAQ--KQYGVADVEVLAAS 329
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 222 RPEWSQK-----VYNAIVKAQW---TRDHRTEAALLSRAGPNKYPHEPAVLTKY 270
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 330 M-----NGTFEBKDKRALPDENTFFPHGASYPSSAVWVLTQLRKGMIN 377
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 271 LAPAAEDRAGYIASGIRIQWDEKRI--DFQY-----PFSYTB--LVRLKTTLL- 320
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 EFKPDWVYLDTAKNVYRPDIYLAALKELV 406
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 321 --EGDNTFLSGLDPAV-----AARDLV 340
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9

```

US-10-282-122A-49957
; Sequence 49957, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

```



```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48182
; LENGTH: 391
; TYPE: PR1
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48182

```

```

Query Match      8.9%; Score 217.5; DB 15; Length 391;
Best Local Similarity 23.3%; Pred. No. 5,7e-11;
Matches 95; Conservative 6; Mismatches 130; Indels 115; Gaps 16;

```

```

QY 9 SKLLLTLSASLAWG---LTTA-----PVGAVGKLEKEDLKFGIKLTDMAPLAV 57
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2 SREWMKLSLFTATGAPLISMANAAMAEPA-----PVRIGYLPITDAAPLAV 52

QY 58 AAKEKGFEDGFLV-QLEAQAANKVMVDRVNGELDGSMLAPAPLAASGFEKTDIEV 116
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 53 AHNNGFPASGELVVEOPKLRSAQVLEAFLSQVNVVHLISMTLMAR--YGSRAPAKV 110

QY 117 PPSMGFNAGNATVSNIEIWHQMKPNIPLEGSKPVHPRIKADYLKPVVEKYKAEGKPFNMAMT 176
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 111 VAMNHVNGSGLTVAPI-----DSLRLDGGKTV-----AVP 141

QY 177 FPAASHNITKRYWLAAGINPGYISPPDISGOIGADALISVTPPQMPSTLEAGTTFGV 236
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 142 FWSJHNVVLQDMLRSGGLVPLVKRRGTPANAEVN-----LVVMAPSDMLPALAARQIAGY 197

QY 237 CVGEPNQOAVFGIGVPTVDEELMKDPEKVFGLTKQ-----MAEKYPTIYLAAYK 289
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 198 IYAEPPNATBELMKIGLIFETGDVMKNHACCVFVHEBDLTLORPMSQK-----VVA 250

QY 290 ALIRAIWLDADNNKREALEMIAQOYGVADVEVLAASMGTFEYKED-----DKRAL 344
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 251 AIVKADLMA-----RSHPQETAOUL-----SKDTHHYSPTTLASLDRLV 291

QY 345 PD---FNTFPRHGA-----SYVSYSAAWYLTQLR 372
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 292 PSASLADTYRASGAIRHADWHAKRIDFOYPFPSTYEAAL--VQRLKR 336

```

```

RESULT 12
US-10-282-122A-72603
; Sequence 72603, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau

```

```

; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72603
; LENGTH: 342
; TYPE: PR1
; ORGANISM: Streptococcus mutans
US-10-282-122A-72603

```

```

Query Match      5.9%; Score 144; DB 15; Length 342;
Best Local Similarity 21.4%; Pred. No. 0.00028;
Matches 83; Conservative 59; Mismatches 145; Indels 100; Gaps 15;

```

```

QY 12 LLLTSLASLAWGLTAPVAVGKLEKEDLKFGIKLTDMAPLAVAAEKGFEDDEGL 69
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 15 ILFISVSLVACSKSTSTSKNYNASEKDSYILKISENSDLGAPQOLAIIEGFFDDVL 74

QY 70 FVQLEAQAANKV-----MDRVNGELDGSMLAPA---PLAASVFGTKADIEVPE 118
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 75 -----KRYVYKIGQDTSNLDALNAGKIDSNSLMASTIOPLA----- 111

QY 119 SMGFNGNATVSNIEIWHQMKPNIPLEGSKPVHPRIKADYLKPVVEKYKAEGKPFNMAMTP 178
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 112 ---NGAKUKITGTL-HTSCQLILTKNGK-----ISAAELKKKKIGTVAV 152

QY 179 AASHNITKRYWLAAGINPGYISPPDISGOIGADALISVTPPQMPSTLEAGTTFGV 238
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 153 AGSPALFAKRYVLAKSKGLK-----VSEDEKG-DVSFVTVQSDQLGVLIDGAEVDAL 202

QY 239 GEPNQOAVFGIGVPTVDEELMKDPEK---VFQVTKQMAEKYPTNYLATKALIRAA 295
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 203 GDP-DTEVILAKQYGFRTLANSTDGFKRKYCVCAVVSNDYIKKIPPAVAAKTTLAMQKAA 261

QY 296 IWLADNNKREALEMIAQOYGVADVEVLAASMG--TFEYKEDKRALPDPTTFPRHG 354
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 262 MNV-----QGHKEETVDIQLNQYVAGSKDSNLTSLNSTYFK----- 298

QY 355 ASYPSYSAAV---WYLTQLRWGMIN 377

```

Db 299 ---PSYSGAYDSFDTVASDLRKIGILS 322

RESULT 13

US-10-282-122A-47676
; Sequence 47676, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Hasebeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreysch, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47676
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-47676

Query Match 5.7%; Score 140; DB 15; Length 192;
Best Local Similarity 26.8%; Pred. No. 0.00027;
Matches 42; Conservative 32; Mismatches 51; Indels 32; Gaps 4;

QY 42 LKGGFKLTDMALVAAEKGFPEDGGLPV-QLEAQAANKVMDRVNNGELDGSHTLA 100
DB 37 VRIIGYLPITDAALLVAHNNGVPASGELTVEQPKLLRSWAQVLEAFSGQVNVVHLLSPM 96
QY 101 PLAASVGFQTKADIEVPFSGFNALITVSNIEIWHQMKPIIPLEGSKPVHPRIKADYLV 160
DB 97 TLMAR--YSGRAAKVVAAMHVNNGSLTVAAPDI-----DSLRLGSKTV----- 138
QY 161 VEKYKAGKRFNNAMTFPAGSHNIKLRWYLAAGINP 197
DB 139 -----ALPFWYSIHNVVLQDMLEQGLVP 162

RESULT 14
US-10-156-761-8511

; Sequence 8511, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8511
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8511

Query Match 4.8%; Score 116.5; DB 14; Length 883;
Best Local Similarity 21.3%; Pred. No. 0.41;
Matches 114; Conservative 62; Mismatches 147; Indels 213; Gaps 35;

QY 6 RSSSKKLLITTSASLAVNGLTAPDVGA-----VGKLEKEDLKFGFKLTDMAP---L 55
DB 15 RSLGRSLALTWAVVGAAGLALSPASADSGPTTADPTEAHGKGEV--YTQSAAGADF 72
QY 56 AVAAKGGFPEDEG-FVQLEAQAANKVMDRVNNGELDGSHTLAAPLAASGFQTKADI 114
DB 73 AELKATGF--DPGLDFDSLSR-----LSAATGRADVDV 105
QY 115 E-----VP-----FSM-GFNALITVSNIEI-----W--HQMKPIIPLEGSKPVHP 151
DB 106 RWTGEVVEBKSGPTTFSTIGDNGFRLWDKLTIDHWVDMDRQTSQPVDTLAGR-AVD 164
QY 152 IKADYLVKRVVEKYKAGKRFNNAMTFPAGSHNIKLRWYLAAGINPGYSPQ----- 204
DB 165 IKEYFEFH-----GGSNHLHR--WTPPGSKS---AVQSAFRLPD 201
QY 205 --DISGQAGDAL-----LSVPPRPQMP--STLEAGTIFGVC 237
DB 202 GFDVNGALSAIVLGDGRVRLDFQALASAPATVLDHLITVSGKWPUSVKT----- 254
QY 238 VGEPMNQOAVFKGIGVPIYTD-----ELMKD-----TPEKVFQVTRQWARK 279
DB 255 --DPSDARALLVGLGEPIVGNKAGKAGVAVTYDGKALADTSGTPYKAF-----WSSG 307
QY 280 YFN--TYLAVTKALIRAILWLDADNNKRNKEAIEMLAQOYVADDEVILA-ASNNGTPEY 336
DB 308 -FNKSTYELRTE-----MADQVGRK-----ALPEYRPPQLTRKAMQNLNGTWQF 351
QY 337 -----EKDDKRALPDNTPFRPHGASVP-----SYSAAVVYVQLRWGMINERK 380
DB 332 AGAKAGGEPPRTGRTLA-----RIIVFPVBSQSGIERHEDRMV-----RTFTV----- 398
QY 381 PDNWTYLDTRAKV-----YRPDIYLAAKEVLAEGKAKAEDFPADT--SIKPSQ 426
DB 399 PADWVKVGHGKRLRLNFGAVNDWQSEVYVNGTK--VAEHKGGYDKFSDADVTDALKPER 452

RESULT 15
US-10-450-763-36085
; Sequence 36085, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hybee, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO: 36085
LENGTH: 1355
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (287)...(307)
OTHER INFORMATION: Elongation factor 1 beta/beta'/delta chain proteins domain
OTHER INFORMATION: Identified by eMatrix, accession number BL008248, p-value=8.419e-
US-10-450-763-36085

Query Match 4.6%; Score 113; DB 18; Length 1355;

Best Local Similarity 21.6%; Pred. No. 1.7; 196; Indels 152; Gaps 25;

Matches 110; Conservative 51; Mismatches 196; Indels 152; Gaps 25;

QY 2 KTIIRSSSKLLITLSA-----SLAVGLTIAPDVAVGKL--EKEDLKEFTK 48
DB RTDVASESEKGLTRSAQETVKHSDLFSSSSPMDKGTKPTKTVLSLPDEED-----K 843
QY 49 LIDMAPIAAAEK-----GFPEDEGLF-----VLEAQNMYK 81
DB 844 MEDSKPLSQAPQKEVGCDPDAPKSTGVFQDEELFSHKLQKNDPPVDLFAGTKTK 903
QY 82 VMDRVVN--GELDSHMLAPA--PLASVGGTADIEV--PFSMGFNGNAITVSEIM 134
DB 904 LLEPIVGSLEFGDEDDDLSSAKSQPLVQEKRVKKDHSVDSFKQKHPESTIQSKKKG 963
QY 135 HQMKPNIPLEGKSPVPIKADYLPVVEKYKAEKPFNNAMTFPPAGSHNIKLRWYLAAG 194
DB 964 HMKKPEPT-QANLAINPAA--LLPTAASQISEVKVLPPELAFPSEH--RRSHGLESP 1017
QY 195 INPGYSSPPQDISGQIG-----ADALISTPP-----POMESTLEAGTIFGYCVGB 240
DB 1018 VLPG-----SGEAGVSFDLPQAQDTLHSANKSRVKRGKRRPQTRAAARLAA----- 1064
QY 241 PMNQAVFPGIGVPIVTDELMWDTEKVFQVTKQWAEK--YPTNYLAVTKALIRAIWL 298
DB 1065 --QESSEAEADMSP-----RPIAQADGALSPNGH---RPQLRA----- 1099
QY 299 DADNNKRRKAIEMLAQK-----QYVADVEVLAASNGTFEYKODKRALPDNTFFRHG 354
DB 1100 -ASGEDSTEALAAAPAEWGGVPPGVDSPPAKSL-GHSRGAD---LFDSDGDIFFSTG 1153
QY 355 ASYPSYSSAVVLTQLRWGMINEFEKPDWYLDPTANVTRPDIIYLAARELVAEGKAE 414
DB 1154 TGSQS-----VERTKPK--AKIAENPANPV-----GGRKAKSP 1184
QY 415 DEPADTSIKPSONFIDKVPFDANKPNDY 443
DB 1185 MFPALGEASSDDDLFOSAKPKPAKKTNP 1213

Search completed: September 16, 2005, 01:02:42

Job time : 178 secs

This Page Blank (uspto)